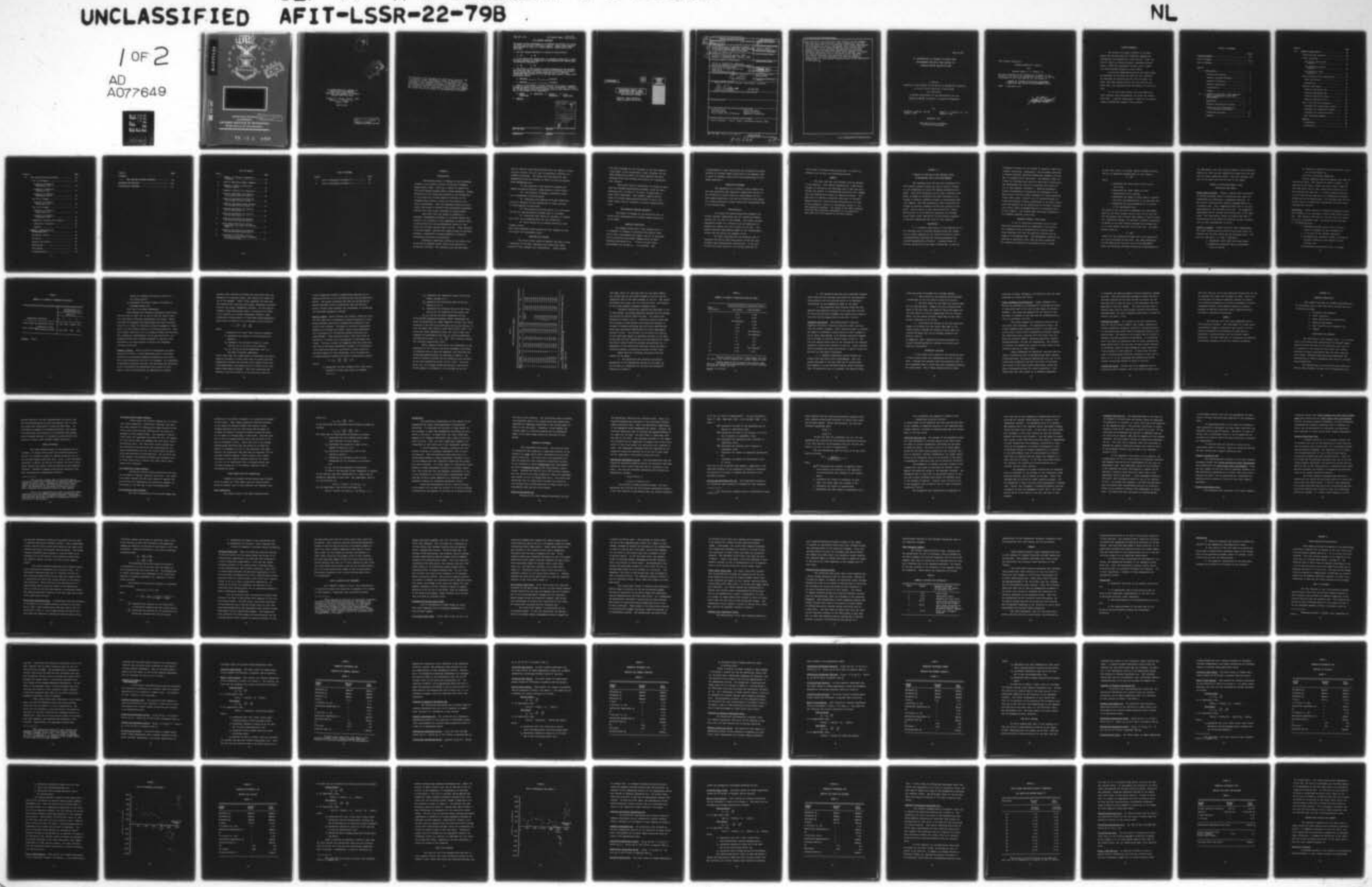
AD-A077 649 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL--ETC F/G 5/1
AN INVESTIGATION OF CHANGES IN DIRECT LABOR REQUIREMENTS RESULT--ETC(U
SEP 79 M W CROZIER, E J MCGANN
UNCLASSIFIED AFIT-LSSR-22-79B NL

1 OF 2
AD
A077649



LEVEL II

3

B.S.

UNITED STATES AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY
Wright-Patterson Air Force Base, Ohio

3

AN INVESTIGATION OF CHANGES IN DIRECT LABOR REQUIREMENTS RESULTING FROM CHANGES IN AIRCRAFT ENGINE PRODUCTION RATE

Michael W. Crozier, Captain, USAF
Edward J. J. McGann, Jr., Captain, USAF

LSSR 22-79B

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deleterious information are contained therein. Furthermore, the views expressed in the document are those of the author(s) and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the Air Training Command, the United States Air Force, or the Department of Defense.

USAF SCN 75-20B AFIT Control Number LSSR 22-79B

AFIT RESEARCH ASSESSMENT

The purpose of this questionnaire is to determine the potential for current and future applications of AFIT thesis research. Please return completed questionnaires to: AFIT/ LSH (Thesis Feedback), Wright-Patterson AFB, Ohio 45433.

1. Did this research contribute to a current Air Force project?

a. Yes b. No

2. Do you believe this research topic is significant enough that it would have been researched (or contracted) by your organization or another agency if AFIT had not researched it?

a. Yes b. No

3. The benefits of AFIT research can often be expressed by the equivalent value that your agency received by virtue of AFIT performing the research. Can you estimate what this research would have cost if it had been accomplished under contract or if it had been done in-house in terms of manpower and/or dollars?

a. Man-years ________ $ ________ (Contract).

b. Man-years ________ $ ________ (In-house).

4. Often it is not possible to attach equivalent dollar values to research, although the results of the research may, in fact, be important. Whether or not you were able to establish an equivalent value for this research (3 above), what is your estimate of its significance?

a. Highly Significant b. Significant c. Slightly Significant d. Of No Significance

5. Comments:

Accession For	
NTIS GRA&I	☑
DDC TAB	☐
Unannounced	☐
Justification ________	
By ________	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	

________________ Name and Grade ________________ Position

________________ Organization ________________ Location

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE. $300

BUSINESS REPLY MAIL
FIRST CLASS PERMIT NO. 73236 WASHINGTON D.C.

POSTAGE WILL BE PAID BY ADDRESSEE

AFIT/LSH (Thesis Feedback)
Wright-Patterson AFB OH 45433

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

14

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFIT-LSSR-22-79B	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER 9
4. TITLE (and Subtitle) 6 AN INVESTIGATION OF CHANGES IN DIRECT LABOR REQUIREMENTS RESULTING FROM CHANGES IN AIRCRAFT ENGINE PRODUCTION RATE.		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 Michael W. Crozier, Captain, USAF; Edward J.J. McGann, Jr, Captain, USAF		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS School of Systems and Logistics Air Force Institute of Technology, WPAFB, OH		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Department of Communication and Humanities AFIT/LSH WPAFB OH 45433	11	12. REPORT DATE September 1979
		13. NUMBER OF PAGES 113
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 12 128		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
JOSEPH P. HIPPS, Major, USAF Director of Information	18 SEP 1979	
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Learning Curve; Cost Estimating; Direct Labor Cost Estimating	Production Rate; Engine Production; Cumulative Production	
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Thesis Chairman: John E. Engel, Lieutenant Colonel, USAF		

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

012250 LB

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

The addition of the production rate variable to the standard learning curve model has been studied extensively, and has been validated for use in estimating direct labor requirements for airframes and avionics equipment. This research set out to determine if the technique was applicable to aircraft engines. Three aircraft engine production histories were evaluated under the model. Of these three engines, one engine, the Pratt and Whitney F-100, was found to fit the model. The authors concluded that the model could be useful in some engine production programs, but that the specific program must be individually evaluated prior to the application of the model.

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

LSSR 22-79B

AN INVESTIGATION OF CHANGES IN DIRECT LABOR REQUIREMENTS RESULTING FROM CHANGES IN AIRCRAFT ENGINE PRODUCTION RATE

A Thesis

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Logistics Management

By

Michael W. Crozier, BS, MA
Captain, USAF

Edward J. J. McGann, Jr., BS
Captain, USAF

September 1979

This thesis, written by

Captain Michael W. Crozier

and

Captain Edward J. J. McGann, Jr.

has been accepted by the undersigned on behalf of the faculty of the School of Systems and Logistics in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN LOGISTICS MANAGEMENT
(CONTRACTING AND ACQUISITION MANAGEMENT MAJOR)

DATE: 7 September 1979

John E. Engel
COMMITTEE CHAIRMAN

ACKNOWLEDGEMENTS

The authors are deeply indebted to the many people who provided help and thoughtful suggestions throughtout the preparation of this thesis. First, we wish to thank our thesis advisors, Lieutenant Colonel Larry L. Smith and Lieutenant Colonel John E. Engel, for their valuable guidance in this effort.

Our sincere appreciation goes to the many people who assisted in providing the data for this research: Art Mohliman and Richard Ante for the J-79 data, Jim Snow, John Baum, Konnie Wilcox, and Jim Link for the TF-41 data, and Doug Taylor and Ron Runkle for the F-100 data.

To our wives Pegi Crozier and Lola McGann for their patience and understanding, we extend our eternal gratitude. A special thanks goes to Pegi for her expert, timely, professional typing of this thesis.

TABLE OF CONTENTS

Chapter		Page

Chapter		Page

APPENDIX

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The American public is taking an ever-increasing interest in the manner in which the Federal Government spends public funds. There are a number of reasons for this trend towards greater public fiscal awareness. Rising inflation and the continuing decrease in the buying power of the dollar have focused people's attention on the increased taxes needed to carry out the numerous functions of government. Also the widespread publicity given recent discoveries of scandal within government purchasing agencies has fueled public mistrust of big government.

Each year, when the budget is presented to Congress, considerable attention is directed towards that portion to be allocated to the Department of Defense (DOD), a portion which has been substantially reduced in real terms over the past several years (12:273). This reduction in the DOD budget, combined with higher prices for military hardware, necessitates increased emphasis on improved purchasing techniques and procedures.

Learning or experience curves are one method used by the DOD to estimate direct labor hours required in production (13:2D25). The learning curve was developed

from an observed relationship whereas when the number of units produced doubles, the unit cost (in manhours) decreases by a constant percentage. The theory was developed with some assumptions about the industry to which it could be applied. These assumptions are:

(1) The production item should be sizeable and complex and should require a large amount of direct labor.

(2) The majority of assembly operations should not be mechanized or machine-paced.

(3) Learning curves applied from past experience should be adjusted for any differences in items, process, or other aspects of production.

(4) The production process should be a continuous one and the item and product changes kept to a minimum.

(5) Historical data should be available to compute the curve, since estimated data has low reliability.

(6) There should be no external production rate changes (2:231).

This final assumption about production rate changes has been the subject of considerable study.

LIMITING THE PROBLEM

The focus of much previous research has been to seek a variation to the basic learning curve model which takes production rate changes into consideration. These studies

and their findings will be discussed in the next chapter. A key result of the studies was a model developed for airframe production by Smith (11) which improved the basic learning curve model through the addition of a production rate variable.

This model should be applicable to industries other than the airframe manufacturing industry, since the basic learning curve, on which it is based, is applicable to these other industries. This research applied Smith's model to aircraft engine production data to see if the basic theory could be extended to aircraft engine production.

THE RESEARCH PROBLEM STATEMENT

The effect of changes in the production rate on direct labor requirements for aircraft engine production is not known.

RESEARCH OBJECTIVES

The primary objectives of this research are (1) to determine if changes in production rate affect total direct labor requirements per engine, and (2) to determine how the production rate model compares with the basic learning curve model as a predictor of direct labor requirements for aircraft engine production. The

accomplishment of these objectives will determine the applicability to aircraft engine production of the approach used by Smith to estimate direct labor hour requirements in airframe production.

RESEARCH HYPOTHESES

The hypotheses to be tested in this research are that the production rate explains a significant portion of the variation in direct labor hour requirements for aircraft engine production, and that the production rate model is a better predictor of direct labor hour requirements than is the reduced model.

JUSTIFICATION

In a major DOD system acquisition program, the initial contract negotiations set up a tentative monthly schedule for the life of the program. The formal contract agreements between the contractor and the Department of Defense are usually only for a period of one year. Subsequent delivery requirements depend on the funding appropriated for the project by Congress (11:2).

The basic learning curve is currently used to determine the direct labor hour portion of the new cost estimates used during annual negotiations. If the production rate model proves to be a better predictor of

labor hours in aircraft engine production, it would be a valuable tool in this cost estimating process.

SUMMARY

The topic has thus been narrowed from the general to the specific: from federal expenditures, to DOD expenditures, to the use of the learning curve in DOD cost estimating, and finally to incorporating production rate changes into the learning curve model for improved accuracy. The application of Smith's production rate model to aircraft engine production, which has been identified as the focus of this research, could be of substantial benefit to the DOD acquisition process. The theory of the basic learning curve and the historical development of the production rate model will be presented in the next chapter.

CHAPTER II

A HISTORY OF THE USE OF THE LEARNING CURVE IN ESTIMATING LABOR HOUR REQUIREMENTS

The learning curve has been used extensively in the aircraft industry during the last 30 years to assist in cost estimating for major DOD weapons acquisition programs. Since the introduction of the basic learning curve model, a number of variations have been developed in an attempt to achieve a greater accuracy in predicting actual cost figures. The most promising of these variations developed in previous research has been the production rate model. Smith's study in this area resulted in a model which proved to be an effective predictor for airframe direct labor costs; replication of this model as applied to aircraft engines was the focus of this research.

BEGINNINGS

T. P. Wright's 1936 article on the application of the learning curve to aircraft manufacturing cost estimation is widely regarded as the initial substantive effort in mathematically modeling the learning phenomenon for aircraft manufacturing (13:2D26). Increased rates of aircraft production at the onset of World War II drew US

Government attention to the concept of improved output with constant facilities; consequently, the Government sponsored a Stanford Research Institute statistical analysis of World War II airframe direct labor data. The Stanford study confirmed the learning curve effect on World War II production, and demonstrated the value of a learning curve model for use in cost analysis (13:2D26-7). Numerous subsequent studies concentrated on validation and refinement of the basic model to achieve better estimates of future projected direct labor costs. Specific foci of various studies included the addition of more coefficients to the basic linear model and analysis of the reliability of the basic learning curve model; these investigations will be addressed later in this chapter. All studies dealt with some variation of the standard learning curve model.

STANDARD LEARNING CURVE MODEL

It can be intuitively discerned that for labor production processes which are repetitious, each successive equivalent unit of production will require fewer direct manhours, and that the manhours required will decrease at a decreasing rate. This phenomenon, known as the learning or experience curve, has two basic variations. The variation validated by the Stanford study is known

as the "unit curve" or "Boeing" theory (13:2D28; 10:273), and can be expressed mathematically by the formula:

$$Y = AX^B$$

where:

Y represents the direct labor hours for the "xth" unit;

X represents the total number of units manufactured in the process;

A represents the number of labor hours to produce the first unit manufactured in the process; and

B represents the slope parameter or a function of the improvement rate.

The slope of the curve can be expressed as a percentage, which is the ratio between the per unit cost at any unit and the percent cost at double that number of units (2:199). The "cumulative average" or "Northrop" variation (described by Wright in his 1936 article) measures the average cost for x units rather than cost for the xth unit. Its mathematical form is:

$$\bar{Y} = AX^B$$

"where $\bar{Y}$ is the cumulative average cost of all production up to and including the xth unit. The other parameters are the same as for the unit curve theory [13:2D29]." While the Boeing and Northrop models can be manipulated in

the same manner, the user should be aware of the difference between the unit cost and cumulative average unit cost measured by these respective models. The unit learning curve will be the model used for the rest of this paper.

HISTORY OF THE DEVELOPMENT OF THE PRODUCTION RATE MODEL

Robert Blair Ilderton. Ilderton compiled data on methods of fitting learning curves to production data using the assumptions and techniques of regression analysis. He was chiefly concerned with computer programs written to accomplish this. Also, suggestions were made on possible improvements to the learning curve model. One suggested improvement was to consider more factors than in the basic model. The first factor listed by Ilderton was that of consideration of the effect of different levels of production on labor requirements (5:69-71).

Gordon J. Johnson. Johnson predicted labor requirements for rocket motors using an additive form model which considered both the rate effect and the learning effect. The model he used was $y = A + BX_1 + CX_2^Z$ where:

y represents direct labor hours per month;

X_1 represents production rate in equivalent units per month;

X_2 represents cumulative units produced as of the end of each month; and

A, B, C and Z are model parameters.

Johnson regressed this model against four sets of rocket motor data. His results are shown in Table 1. As depicted in the table, Johnson had good results (high R^2) with data sets 1 and 4, fair results with data set 2 and poor results with data set 3. Johnson explained data set 3's poor results as being due to an inadequate accounting system used by the manufacturer. He concluded that the production rate is a significant determinant of direct labor requirements (6:25-41).

<u>Joseph Noah</u>. Smith reported on research performed by Noah, with emphasis on his findings on direct labor hours (11:31-35). Noah analyzed cost data to find the effect of production rate on airframe costs. His model for the data was $y = e^A \cdot X_1^B \cdot X_2^C \cdot X_3^D$ where:

y represents average direct labor hours per pound of airframe for each airframe lot;

e is the base of the natural logarithm;

X_1 represents the cumulative volume in pounds of aircraft produced by the midpoint of each airframe lot;

X_2 represents the production rate in average

TABLE 1

SUMMARY OF JOHNSON'S REGRESSION ANALYSIS

Regression Variables	Coefficients of Determination (R^2) Data Set 1	2	3	4
Labor Hours vs Cumulative Units	.753	.395	.00678	.763
Labor Hours vs { Cumulative Units & Production Rate	.932	.808	.308	.927

Source: (6:34).

pounds of airframe delivered per month for the entire period;

X_3 represents the annual volume of aircraft in airframe pounds; and

A, B, C and D are model parameters.

Noah averaged the estimated regression coefficients from two sets of data, one on the F-4 and the other on the A-7, and tried to develop a generalized cost model. Smith felt that this approach was questionable and that the model needed to be tested on additional aircraft programs to determine if it did actually serve as an accurate predictor. Also, Smith stated that while the lot average airframe delivery rate was a practical representation of the production rate, the average delivery rate variable appeared to lag the actual expenditure of hours required to produce the airframes delivered (11:31-35).

Joseph A. Orsini. Orsini (9:57-80) investigated the applicability of Johnson's three-dimensional model to airframe production by performing regression analysis on the direct labor hours associated with the C-141 program. He established 24 quarterly production periods, and calculated the equivalent units of production and associated labor hours for each period. The C-141 production run experienced a typical

quantity rate fluctuation whereby the production rate rose gradually to a maximum output, then tapered off toward the end of the program. Orsini first regressed the data with the standard unit learning curve model, employing cumulative units as the single independent variable. He next applied Johnson's three-dimensional model to determine the effect of the addition of a second independent variable, then introduced a multiplicative or log-linear three-dimensional model which is stated in its curvilinear form as follows:

$$Y = e^{B_0} \cdot X_1^{B_1} \cdot X_2^{B_2}$$

where:

Y represents the direct labor hours per quarter;

X_1 represents the number of units produced per quarter;

X_2 represents the cumulative number of units produced as of the end of each quarter;

B_0, B_1, and B_2 are model parameters, and e is the base of natural logarithms.

Orsini found that the coefficients of correlation and determination decreased in every case where the production rate independent variable was excluded, which implied that "the three-dimensional cost models provide a better fit for the sample data analyzed [9:70]." His first conclusion was therefore that the inclusion of the production rate as a

second independent variable significantly improved the explanatory ability of both the additive and the multiplicative models. His second conclusion was that the multiplicative models performed better as predictors than the additive models "because they eliminate the requirement of estimating the additional parameter Z [9:71]."

Larry L. Smith. Smith directed his research toward the problem that there was no generally accepted technique for estimating airframe costs which systematically employed a production rate variable. He sought to develop a generalized approach to building cost models, where "model parameters can be tailored to a continuing airframe production program [11:56]." These tailored models could then serve as predictors of direct labor requirements of additional airframes. Previous studies had suggested that the production rate was an important independent variable in the formation of an effective predictor, and Smith's initial tests had provided support for this hypothesis, so he developed the following modified version of Orsini's multiplicative model:

$$Y_i = B_0 \cdot X_{1i}^{B_1} \cdot X_{2i}^{B_2} \cdot 10^{e_i}$$

where:

Y_i represents the unit average direct labor hours required to output each pound of airframe in lot i;

X_{1i} represents the cumulative output of all airframes through lot i;

X_{2i} represents the production rate of the airframes in lot i;

e_i represents the variation which is left unexplained by the variables in the model; and

B_0, B_1, and B_2 are model parameters (11:43).

To facilitate regression analysis, Smith transformed the model by extracting the common logarithm of each term, which resulted in the form $\text{Log } Y_i = \text{Log } B_0 + B_1 \text{ Log } X_{1i} + B_2 \text{ Log } X_{2i} + e_i$, with the model thus linear in each term (11:19; 3:45). He called his three-dimensional model the "full" model, and compared its performance with that of the two-dimensional multiplicative model, $Y_i = B_0 \cdot X_{1i}^{B_1} \cdot 10^{e_i}$, which he called the "reduced" model (11:69; 3:19).

Smith compiled direct labor hour statistics on the production of F-4, KC-135, and F-102 airframes, and formed 16 different sets of data for analysis, with each data set containing characteristics as indicated in Table 2. He used two different proxies for the production rate variable in his three-dimensional model (11:13; 3:19). The first proxy, the "lot average manufacturing rate", was defined as the number of airframes in a lot divided by the lot

TABLE 2

SMITH'S REGRESSION MODEL SUMMARY

Cost Model	Airframe Model	Cases	Level	Rate	R^2 Full	R^2 Reduced	B_0	B_1	B_2
1	F-4A-F	57	Total	Del	0.078	0.928	masked[a]	-0.261	-0.169
2	F-4B-F	55	Total	Manu	0.973	0.904	"	-0.246	-0.183
3	F-4B-F	55	Total	Del	0.966	0.904	"	-0.257	-0.161
4	F-4B-F	42	Total	Manu	0.853	0.585	"	-0.230	-0.157
5	F-4B-F	42	Total	Del	0.820	0.585	"	-0.229	-0.136
6	F-4B-F	42	Fabri	Manu	0.889	0.618	6.328	-0.221	-0.148
7	F-4B-F	42	Fabri	Del	0.851	0.618	7.601	-0.219	-0.127
8	F-4B-F	42	Assem	Manu	0.744	0.658	9.016	-0.279	-0.112
9	F-4B-F	42	Assem	Del	0.733	0.658	10.400	-0.278	-0.097
10	F-102A	50	Total	Del	0.979	0.961	38.371	-0.299	-0.158
11	F-102A	42	Total	Del	0.979	0.959	47.290	-0.344	-0.144
12	KC-135A	96	Total	Del	0.958	0.971	13.133	-0.453	-0.164
13	KC-135A	7	Fabri	Manu	0.974	0.903	0.674	-0.165	-0.305
14	KC-135A	7	Fabri	Del	0.971	0.903	1.123	-0.233	-0.222
15	KC-135A	7	Assem	Manu	0.994	0.964	13.338	-0.608	-0.361
16	KC-135A	7	Assem	Del	0.992	0.964	7.303	-0.527	-0.263

[a]Data considered proprietary by the manufacturer, and therefore masked in the published version of Smith's research (11:65).

<u>Source</u>: (11:71-143).

time span, where lot time span was the time span between the release date of the first airframe in the lot and the completion date of the last airframe in the lot. The second production rate proxy, the "lot delivery rate", was defined as the actual monthly airframe acceptance rate.

Extensive statistical analysis was performed on the 16 data groups, including regression analysis for all groups, and predictive ability tests for most. The results are tabulated in Tables 2 and 3. Predictive ability tests were not performed on data groups (test situation numbers) 13 through 16 because each group had only seven observations, which were considered insufficient to achieve meaningful results (11:131). The accuracy of the predictive ability of each model was tested by omitting a portion of the downstream data and then regressing the model against the remaining data to predict what the omitted values should be. A comparison of the omitted and predicted values provided a subjective measure of predictive ability (11:56; 3:20).

Smith drew the following conclusions from the results of his tests:

1. The production rate variable contributed importantly to the explanatory ability of the model, as can be seen by a comparison of the full and reduced R^2 statistics in Table 2.

TABLE 3

SUMMARY OF SMITH'S PREDICTIVE ABILITY TESTS

Test Situation #	Percent Deviation of Predicted Value	
	Full Model	Reduced Model
1	-2.63	14.50
2	2.19	13.60
3	Not Reported	13.60
4	1.75	5.26
5	3.07	5.26
6	-7.84	Not Reported
7	"a"	Not Reported
8	"b"	1.07
9	"b"	1.07
10	-1.05	5.61
11	3.51	Not Reported
12	2.20	-3.30

[a] Smith reported the deviation was larger than that for test 6, but did not indicate the specific value (11:96).

[b] Smith reported the deviations were smaller than those of the respective reduced models, but did not indicate the specific values (11:101).

Source: (11:71-125)

2. The manufacturing rate proxy performed slightly better than did the delivery rate proxy for the production rate variable, and both proxies proved to be important contributors to the explanatory power of the model.

3. The production rate variable stabilized and improved the predictive ability of the cost model for the F-4 and F-102 aircraft programs, and results were inconclusive with KC-135 program data (11:142-146; 3:21-24).

Congleton and Kinton. Replicating Smith's approach, Congleton and Kinton evaluated historical data collected from the T-38/F-5 program. Their objectives were to identify the impact of externally caused production rate changes on direct labor requirements in an ongoing production program, and to attempt to validate Smith's model (3:26-27). The full and reduced models used in the testing and the research hypotheses tested were identical to Smith's. Their research effort validated the predictive ability of Smith's model and provided the following conclusions:

1. A negative correlation exists between production rate and direct labor requirements. This conclusion was based on the fact that the R^2 coefficients were negative in every situation tested, which indicated that "if production rate is increased, the required labor

hours per pound of airframe will decrease [3:91]."

2. Both production rate proxies were important contributors to the full model's predictive ability.

3. The full model fit the data more closely than did the reduced model. Support for this conclusion was based on the fact that the actual R^2 value was higher for the full model in 29 of 30 comparisons with the reduced model, and the single contrary result was at less than 1 percent difference in R^2 values.

4. The predictive ability of the full model was judged to be better and more stable than that of the reduced model, although the results on which this conclusion was based were not as strongly supportive as Smith's (3:91-5).

In summation, then, Congleton and Kinton succeeded in closely replicating Smith's results, and further validated his conclusions.

DISSENTING OPINIONS

In the study of the effect of the production rate on direct labor hours, all researchers were not in agreement. Although most studies found the production rate to be an important factor, there were some dissenting opinions and conclusions. Two of these studies were the study

conducted by Large, Hoffmayer, and Kontrovich and the study performed by Alchian and Allen.

Large, Hoffmayer and Kontrovich. Large, Hoffmayer and Kontrovich performed an investigation concerning the effect of production rate on the cost of some types of military hardware. The study was sponsored by the Office of the Secretary of Defense in the hope that an estimating model could be developed (7:49-50).

This study assumed that production rate and unit cost would vary inversely. In conducting the study, the authors examined data pertaining mainly to aircraft production with a brief look at missiles and aircraft engines. The researchers were concerned with four major cost elements for aircraft frames: manufacturing labor, manufacturing materials, tooling, and engineering. The research on missiles and engines was briefer but also looked at the different aspects of costs (7:50).

Large, Hoffmayer and Kontrovich concluded that the effect of production rate on manufacturing labor, manufacturing materials, tooling and engineering could not be predicted with confidence, and that the worth of including this factor in the model seemed questionable during the early development process of a major acquisition. They stated that each case needed to be examined separately

to determine the specific manner in which production changes occurred. This was considered necessary because the way a rate was achieved and the time allowed to achieve the new rate were felt to have as much or more importance than the magnitude of the rate change. The team concluded that a similar model could be used during a contract's production planning phase, but that a separate study would be needed to develop this model (7:51).

Dreyfuss and Large. In a report by the Rand Corporation done for the Air Force in March 1978, Large, working with another researcher, Dreyfuss, arrived at conclusions about the production rate effect which contradicted those of his earlier work. The subject of this report was the effect of extended low-rate production of airframes. In concluding about the effect of production rate on costs, Dreyfuss and Large stated that a very low rate of early output resulted in additional costs and that subsequent higher production rates resulted in cost benefits. In this study the data used in studying labor hours was identical to that used by Smith, so it is not surprising that Dreyfuss and Large reached conclusions similar to Smith's (4:18-21).

Alchian and Allen. Alchian and Allen suggested that as production rates increase, unit costs should increase also.

They felt that the use of less efficient workers and the use of overtime will cause this increase in costs. Their conclusions were not based on empirical studies but rather on a subjective analysis of the work environment (1:19-20). Subsequent studies, which applied rigorous scientific analysis to specific production data, disproved Alchian and Allen's subjective conclusions.

SUMMARY

The learning curve has found widespread use in DOD acquisition efforts. The development of a model which incorporates changes in production rate proved to be an effective predictor for airframes and might also be applicable in related areas, such as aircraft engine production. The next step will be to describe the specific methodology necessary to test the research hypotheses.

CHAPTER III

RESEARCH METHODOLOGY

This chapter outlines the research hypotheses and the procedures used to test them and is divided into six sections as follows:

1. Objectives and approach;
2. Model variables;
3. Model definitions and assumptions;
4. Research hypotheses;
5. Data collection and treatment; and
6. Summary.

OBJECTIVES AND APPROACH

The objectives of this research were: (1) to determine if the direct labor requirements for aircraft engine production were affected by the production rate, and (2) to determine if the production rate model was a better predictor of labor requirements than the basic learning curve model. Meeting these objectives established the applicability of Smith's production rate model to aircraft engine production.

The approach was to collect historical production data on three aircraft engines, the J-79 manufactured by

General Electric, the TF-41 manufactured by Allison, and the F-100 manufactured by Pratt and Whitney, and to evaluate these data using Smith's production rate model. As with Smith's research in airframe production, the production rate model was adjusted to the specific data groups and no attempt was made to formulate a general labor hour model to be used in all aircraft engine production.

MODEL VARIABLES

The three variables used in this analysis were as follows: (1) direct labor hours[1] per aircraft engine produced;[2] (2) cumulative number of engines produced; and (3) the aircraft engine production rate. Since it was considered desirable to improve the ability to predict the direct labor hours required per engine, this was treated as the dependent variable. Cumulative engines produced and production rate were considered the independent variables.

---

[1]For one set of data, that of the F-100 engine, historical records were maintained in standard labor dollars instead of labor hours. Consequently, the labor dollar figure served as a proxy for labor hours for this set of data. An explanation of this proxy is provided in the data collection and treatment section of this chapter.

[2]Due to the complexity and nature of aircraft engine production, the only feasible measure of production accomplishment was the number of engines completed. Consequently, each aircraft engine was treated as an equivalent unit of production in the research.

The Direct Labor Hours Variable

Direct labor for aircraft engines can be divided into three categories: fabrication, assembly, and test. The labor required for testing was a small portion of the total direct labor and was combined with the labor for assembly by some manufacturers. The dependent variable, labor hours, was expressed three ways for the J-79 engine: as total direct labor, as assembly and test labor, and as fabrication labor. The total labor hours included only that fabrication portion which was accomplished by the engine manufacturer, and did not include labor for parts which were produced by subcontractors. For the TF-41 and F-100 engines, the dependent variable was expressed only as total direct labor, because only total direct labor figures were available.

The Cumulative Output Variable

Records for aircraft engine production were kept of the number of engines completed each month. The cumulative output variable was the total number of engines completed since the beginning of the production program, as calculated on the last day of each accounting month.

The Production Rate Variable

The rate of production of an aircraft engine was

impossible to accurately determine, so a proxy was developed for the model. Since engine production data was recorded by the number of engines finished each month, a daily average production rate for each month was developed by dividing the number of engines finished during each month by the number of working days in that month. This was similar to the lot average delivery rate used by Smith in his work on airframes (11:41). Smith also used a lot average manufacturing rate which was calculated by dividing the number of airframes in a lot by the production time span. However, the production time span was not available for the selected aircraft engines. Smith reported only a small difference in the results obtained using the two models (11:144), so no significant loss of accuracy should have occurred through use of the monthly completion rate in developing the proxy.

MODEL DEFINITION AND ASSUMPTIONS

Chapter II discussed the two models used by Smith which he called the "full model" and the "reduced model". These two models, used in this research, are repeated here:

Model Definition

The reduced model is the basic learning curve

stated as:

$$Y_i = B_0 \cdot X_{1i}^{B_1} \cdot 10^{e_i}.$$

In the full model the production rate variable is added as follows:

$$Y_i = B_0 \cdot X_{1i}^{B_1} \cdot X_{2i}^{B_2} \cdot 10^{e_i}.$$

The terms used in these models are defined as follows:

Y_i represents the unit average direct labor hours required for each engine;

X_{1i} represents the cumulative output of all engines through the i th;

X_{2i} represents the production rate of the engines in group i;

e_i represents the variation which is left unexplained by the variables in the model; and

B_0, B_1, and B_2 are regression coefficients.

In order to use multiple linear regression to analyze the two models, they were transformed to a linear form by taking the logarithm of each term. The logarithmic form of the reduced model is:

$$\text{Log } Y_i = \text{Log } B_0 + B_1 \text{ Log } X_{1i} + e_i$$

and the logarithmic form of the full model is:

$$\text{Log } Y_i = \text{Log } B_0 + B_1 \text{ Log } X_{1i} + B_2 \text{ Log } X_{2i} + e_i .$$

Assumptions

The statistical significance of the results of the regression was tested using appropriate F-distribution statistics. To establish the validity of these tests, it was necessary to make some assumptions concerning the error terms in the model. To begin with, the error terms were assumed to be normally distributed with a mean of zero and equal variance. Also, the error terms were assumed to be independent of each other and of the independent variables.

A third assumption concerns a problem which frequently develops in a multiple regression, that of multicollinearity. Multicollinearity exists when there is a relationship between or among independent variables, which in this research were the cumulative output and the production rate variables. If a high degree of collinearity exists between or among independent variables in multiple regression, the F-test may find the marginal contribution of one or more variables to be statistically insignificant, while they may in fact be good explainers of variation in the dependent variable if considered separately (8:341).

While multicollinearity can be a serious problem if the model is to be used for control, it is not as serious a problem when the purpose of the model is to predict as was

the case in this research. The contribution made by adding the production rate to the reduced model was subjectively evaluated by comparing predictions of the reduced model to those of the full model. Therefore, it was assumed the varying degrees of multicollinearity had no substantial impact on the short-range predictive abilities of the models.

RESEARCH HYPOTHESES

Four hypotheses were tested. The first was that the production rate explained a significant portion of the variation in total direct labor requirements for aircraft engine production. The second and third hypotheses were that the production rate explained a significant portion of the variation in direct fabrication labor requirements and direct assembly and test labor requirements, respectively. (Because the data for the F-100 and TF-41 engines were not divided into these categories, these two hypotheses were tested on the J-79 engine data only.) The fourth hypothesis was that the model which included the production rate variable was a better predictor of total direct labor requirements than was the reduced model.

Research Hypothesis One

Testing of the first research hypothesis involved

two statistical tests and two criterion tests. First, regression analysis was performed on the historical aircraft engine production data. Next, the statistical significance of the regression model parameters was tested. Finally, two criterion tests were used to evaluate the appropriateness of the model for the data. The full model in log-linear form was used in the tests. The dependent variable in the regression analysis was the logarithm of average total direct labor hours (or labor hour proxy) per equivalent unit, measured by month. The independent variables were the logarithm of the cumulative production quantity as of the end of each month and the logarithm of each month's production rate.

Statistical Hypothesis One (A). This hypothesis was that the cumulative production quantity and/or the production rate were related to direct labor hours per unit as indicated in the model. The null and alternate hypotheses were formally established as follows:

$$H_0 : B_1 \text{ and } B_2 = 0$$

$$H_1 : B_1 \neq 0 \text{ and/or } B_2 \neq 0$$

The decision criterion was as follows: The null hypothesis was rejected and the alternate hypothesis accepted if the test statistic F was greater than the critical statistic

F_c at the .05 level of significance[3]. In this statistical test, F = MSR - MSE, MSR = SSR - (p-1) and MSE = SSE - (n-p) where:

MSR represents the mean of the regression sum of squares in logarithmic form;

MSE represents the mean of the residual (or error) sum of squares in logarithmic form;

SSR represents the regression sum of squares in logarithmic form;

SSE represents the residual sum of squares in logarithmic form;

p represents the number of regression parameters; and

n represents the number of observations (8:45, 79, 227-8).

This test of the F statistic was simply a comparison of the explained variance to the unexplained variance to determine the ability of the model to explain the variance of the dependent variable.

<u>Statistical Hypothesis One (B)</u>. This hypothesis focused on the production rate variable to determine if this variable,

---

[3] F_c values were obtained from an F-distribution table (8:807-13).

when combined with the cumulative production quantity variable, explained additional variation in direct labor hours per aircraft engine. Stated statistically, the null and alternate hypotheses were:

$$H_0: B_2 = 0$$

$$H_1: B_2 \neq 0$$

As with the test for hypothesis one (A), the null hypothesis was rejected and the alternate hypothesis accepted if the test statistic F* was greater than the critical statistic F_c at the .05 level of significance.

For this statistical test the value of F* was determined as follows:

$$F^* = \frac{\Delta R^2/g}{(1-R^2)(n-k-1)}$$

where:

ΔR^2 represents the increase in explained variation caused by the addition of the logarithm of the production rate variable to the reduced model;

g represents the number of variables (in this case, one) which cause the increase in R^2;

n represents the number of observations;

k represents the total number of regressors; and

n-k-1 represents the degrees of freedom in the unexplained variation (14:435).

In other words, the F* statistic for this test was the ratio of the increase in explained variance to the remaining unexplained variance, obtained by transitioning from the reduced model to the full model.

<u>Criterion Test One (A)</u>. The aptness of the regression model as a best estimator of aircraft engine direct labor hours was based on certain assumptions concerning the residuals or observed errors. The model could not be deemed appropriate for the data unless assumptions about constant variance of residuals, residual independence, and normal distribution of the residuals could not be rejected on the basis of appropriate tests which will now be described (8:239-50).

The assumption of constant residual variance (homoscedasticity) was tested by plotting the residual values against the predicted values of the dependent variable. The assumption held if there was no discernible pattern to the plot and if most residuals were within one standard error of the estimate (8:239-40). Another, more objective test of the assumption can be made by use of a rank correlation coefficient (8:112).

The assumption that residuals are independent of

each other and of the independent variables was tested by plotting the residual values against the values of each independent variable. The assumption held if no pattern was detected and if the residual values fluctuated randomly around the base line of zero. In this test, the data points were arranged in time order sequence so that any existing pattern could have been detected. Care was taken to insure that trends in unit data which were the result of engineering or model changes were not mistakenly identified as violations of the assumption of residual independence (11:51). An alternate method of testing independence of residuals is the Durbin-Watson test for autocorrelation in which a test value is computed from the sample and compared to a chart interval of values. If the test value falls within the interval, the condition of autocorrelation can be rejected (14:720; 8:112; 358-61).

The assumption of normal distribution of residuals, which is essential to the accuracy of the F statistic tests described earlier, may be tested in two ways. First, the residuals may be plotted on normal probability paper. For the assumption to hold, the plot should approximate a straight line. Second, the chi-square or Kolmogorov-Smirnov test may be applied to test the assumption (8:112, 506). The first method, which is the faster of the two, was used in this research.

Criterion Test One (B). The appropriateness of the model as an estimator of direct labor was further tested by measuring the proportion of variation in direct labor which was explained by the regression model. The multiple coefficient of determination, known as R^2, was used to conduct this test. R^2 was calculated by subtracting the quotient of SSE/SSTO from one. The error sum of squares, SSE, was the summation of all squared residuals, and was formally defined in statistical hypothesis one (A). The total sum of squares, SSTO, was calculated by summing the squared differences between each observed value of the dependent variable and the mean of the dependent variable (8:77).

In the regression model used in this research, the interpretation of R^2 as a valid measure of explained variation was somewhat obscured by the transformation of the model to logarithmic form. The logarithmic values of the deviation in the dependent variable, when squared to compute the R^2 statistic, yielded a ratio which was valid only to this logarithmic form of the model. A more meaningful statistic, R^2(actual), was developed by Smith for use in his research (11:53). To develop this statistic, the SSTO and SSE values were calculated after transforming observed and predicted values of the dependent variable from logarithmic to actual form. R^2(actual) was then calculated by subtracting SSE

(actual)/SSTO (actual) from one, and represented the variation in actual hours which was explained by the regression model.

An appropriate model for the data would explain a high proportion of variation in direct labor, and would consequently yield a high R^2(actual). Therefore, in this criterion test, an R^2(actual) value of .80 or higher was selected as the level at which the model could not be rejected as inappropriate.

If the model was not rejected by either of the statistical tests or criterion tests, its predictive ability was then tested under research hypothesis four.

Research Hypothesis Two

This hypothesis was identical to the first research hypothesis except that direct fabrication labor requirements were substituted for total direct labor requirements. Since the only change involved the model's dependent variable, the same statistical hypotheses and criterion tests were used to evaluate the model, and the procedures for evaluating the hypothesis were identical to those for the first research hypothesis.

Research Hypothesis Three

This hypothesis was identical to the first research

hypothesis except that direct assembly and test labor requirements were substituted for total direct labor requirements. As with the second research hypothesis, the same statistical hypotheses and criterion tests were used to evaluate the model, and the procedures for evaluating the hypothesis were identical to those for the first research hypothesis.

Research Hypothesis Four

A primary objective of the proposed research was to determine the form of the learning curve which was best able to predict direct labor hours in a continuing aircraft engine production program. The importance of an accurate cost forecasting model for use in DOD aircraft engine acquisition programs has been stressed earlier. Once the full model was successfully developed under one of the preceding research hypotheses, its predictive ability was compared to that of the reduced model, to determine which model was more accurate, and further, to determine the practical value of the model's predictive ability.

The fourth research hypothesis was that for 12 months into the future the full model would be a better predictor than the reduced model. A test of this hypothesis with actual predicted data would have necessitated application of the models over a period of time in a continuing production program. To simulate this situation, the full

and reduced regression models were developed with the last 12 observed data points (months) omitted. Then, using these models, omitted values (which were known but not used in estimating the model coefficients) were predicted. The fourth research hypothesis was accepted or rejected based on an evaluation of the deviation of the predicted values from the observed values, for both the full and the reduced models.

The fourth research hypothesis was evaluated using both a statistical hypothesis and a criterion test. The statistical hypothesis was used to determine whether the full model was significantly better than the reduced model in predicting the labor hour values omitted in the prediction simulation. Where the full model was found to be a significantly better predictor based on the statistical test, a criterion test was then applied to establish whether the improved predictive ability of the full model had a practical significance as well.

<u>Statistical Hypothesis Four</u>. A statistical test was performed to determine if the average absolute deviation of the full model ($|\bar{D}_F|$) was significantly less than that of the reduced model ($|\bar{D}_R|$). The average absolute deviation for each model was computed by taking the absolute value of the

difference between the actual and predicted direct labor hours occuring in each test situation, then separately summing the absolute deviations for each model in all test situations. Stated statistically, the null and alternate hypotheses were:

$$H_0 : |\bar{D}_R| \leq |\bar{D}_F|$$

$$H_1 : |\bar{D}_R| > |\bar{D}_F|$$

The hypothesis was tested using the Student's t distribution, in view of the fact that the number of test situations were few (less than 60). The assumptions of normal distribution and randomness of the deviations, examined in research hypothesis one, remained in effect during this test.

The decision rule using the Student's t statistic was as follows:

$$\text{Reject } H_0 \text{ if } t > t_c(.05)$$

where

$$t = (|\bar{D}_R| - |\bar{D}_F|) / \sqrt{(S_R^2/N) + (S_F^2/N)}$$

and

S_R^2 represents the variance of the distribution of deviations obtained with the reduced model;

S_F^2 represents the variance of the distribution of deviations obtained with the full model;

N represents the number of test situations; and

t_c represents the critical t value obtained from a table of Student's t critical values (14:208-15).

<u>Criterion Test Four</u>. Where the improved predictive ability of the full model over the reduced model was proven to be statistically significant, the model was then subjected to a test of practical significance. This test was deemed necessary because (1) the reduced model, although proven to be a statistically less accurate predictor, could still be sufficiently accurate for practical application, or (2) the full model, although proven to be a statistically better predictor than the reduced model, could still be so inaccurate as to be of no value in practical application. In either of the above two instances, the addition of the production rate variable would not be considered worthwhile from a cost/benefit standpoint.

For the criterion test, the predicted versus actual values of the last 12 data points were compared for the full and reduced models. For each model, the percent deviation of the predicted value from the actual value for each data point was multiplied by the equivalent units produced during that month, and the results were summed. This figure was then divided by the total number of units produced over the 12-month period, which yielded the percent deviation of the

12-month prediction from the actual direct labor costs for the period. This criterion test was considered appropriate, because it was the model's ability to accurately measure a year's (or other contract negotiation time span's) direct labor costs, not the weekly and monthly fluctuations in these costs, which was the primary focus of this research. The full model was judged to have the better practical predictive ability only when (1) it had a smaller percent deviation from actual direct labor costs for the 12-month period, and (2) it was subjectively judged to have a sufficiently low percent deviation so as to prove its practical value as a predictor[4].

DATA COLLECTION AND TREATMENT

As in Smith's research (11:57), the accessibility of data was the basis for selection of programs to be studied in this research. Historical data from three aircraft

---

[4]This subjective criterion for hypothesis acceptance or rejection was arbitrarily selected on the basis of an analysis of the DOD acquisition process. Any improvement in predictive ability, no matter how small, could result in substantial savings for high dollar value programs. However, if the increase in predictive ability were small, the value of the model would be questionable for low dollar value programs. This limitation should be considered prior to application of the model.

engine production programs, the J-79, the TF-41, and the F-100, were obtained. The J-79 data set, consisting of 52 data points, was obtained directly from the manufacturer's computerized records. The TF-41 data set, consisting of 126 data points, was obtained from the Defense Contract Administration System Program Office (DCASPRO) located at the manufacturer's plant. The F-100 data set, consisting of 64 data points, was obtained from the F-100 engine branch of the Propulsion System Program Office (Propulsion SPO) at Wright Patterson AFB. For all three sets of data, each data point consisted of one calendar month's production effort. Each month's figures provided the number of engines (or equivalent units) completed during that month, the average direct labor per engine for those engines completed, and the cumulative number of engines produced, as of the end of the month, from the beginning of the production program. Model variables were derived from this data as discussed below.

<u>Treatment For Direct Labor Hours</u>

Due to differences in format among the three data sets, treatment will be discussed separately for each aircraft engine:

<u>J-79 Direct Labor Data</u>. Direct labor hours for the J-79

production program were computed as actual average direct labor hours per engine for all J-79 engines completed during a particular month (data point). These direct labor hours were provided in two distinct direct labor categories: fabrication and the sum of assembly and test. Three separate models were therefore developed for the J-79 engine. Fabrication direct labor hours and assembly and test direct labor hours were summed to produce total direct labor hours (Y_t), the dependent variable in the first model (Model 1). Fabrication direct labor hours (Y_f) was the dependent variable in the second model (Model 2). Finally, assembly and test direct labor hours (Y_a) was the dependent variable in the third model (Model 3).

TF-41 Direct Labor Data. TF-41 direct labor hour data was available in the same format as the J-79 data except breakout between fabrication and test and assembly was not available. Due to two accounting method changes and one production method change during the production program, data values were transformed to provide data continuity for the model. Two models were developed and are discussed next.

The first model (Model 4) incorporated the two accounting changes. At the 15th data point, the manufacturer's accounting system redesignated indirect support as

a portion of direct labor. The increase in direct labor for all of the remaining data points, as a result of this change, averaged 18 percent as calculated by the manufacturer. In order to provide data continuity, direct labor figures for the first 14 data points were each increased by this 18 percent factor. At the 51st data point, the manufacturer's accounting system redesignated material handlers' labor as a portion of direct labor. The manufacturer calculated the resulting increase in direct labor for the remaining data points to be an average of 3.3 percent. Consequently, the first 50 data points were each increased by this 3.3 percent factor. The dependent variable in Model 4 was the resultant monthly direct labor hour figure for each data point.

The second model (Model 5) reflected the production method change in addition to the two accounting changes. The rate of production of the TF-41 engine was drastically reduced beginning with the 101st data point. As a consequence, the manufacturer initiated a new production method whereby fabrication of engine parts was accomplished twice a year at short intervals. These breaks in the fabrication aspects of the production program introduced "negative learning" to the fabrication process. The manufacturer calculated that

the monthly direct labor hour figures were increased by 21.81 percent as a result of this production method change. Accordingly, all direct labor figures following this change were adjusted by a factor of 1/1.2181 to provide continuity for data before and after the production slowdown. The resultant direct labor hour data was used in Model 5. Model 4 and Model 5 were both analyzed in this research because of lack of substantiation for the manufacturer's 21.81 percent production method change figure.

<u>F-100 Direct Labor Data</u>. The direct labor data obtained from the manufacturer through the Propulsion SPO for the F-100 engine were provided as actual average direct labor costs (dollars) per engine for all engines completed during a particular month (data point); no direct labor <u>hour</u> figures were available. All dollar figures were standardized to January 1976 dollars using a standardization factor peculiar to the manufacturer's plant; consequently, the loss of accuracy in using direct labor dollars in place of direct labor hours for this data set was considered minimal. The standardized dollar figures, a proxy for direct labor hours, were used as the dependent variable in Model 6.

<u>Treatment for Cumulative Output</u>

The manufacturers of all three engines analyzed in

this research maintained accurate records of the number of engines (or equivalent units) which were finished during each calendar month of the production program. Since each calendar month for which the necessary data was available was treated as a data point, the cumulative output variable was computed simply by adding each month's completion figure to the sum of all units completed in the program prior to that month.

Treatment For Production Rate

The production rate proxy used in this research was a daily average production rate for each month, obtained by dividing the number of engines (or equivalent units) finished during each month by the number of working days in that month. The rationale for using this proxy has been described in the model variables section of this chapter. The number of engines finished during each month was obtained directly from the manufacturers' data, as discussed in the preceding paragraph. Each manufacturer's plant operating schedule and holiday schedule were used to calculate the actual number of working days which occurred during each month used as a data point. Care was taken to insure that hours per working day remained constant for all data points in each data set; in those few instances where a working day of shorter duration occurred, that working day was reduced to a

proportionate fraction of the standard working day used on the production program.

Data Treatment Summary

The monthly engine production data, obtained from the manufacturer for the J-79 program, from the DCASPRO for the TF-41 program, and from the Propulsion SPO for the F-100 program, were used to develop the one dependent and two independent variables used in the regression analysis. The nature of the data resulted in the formation of six separate models for regression. These models are summarized as follows:

TABLE 4

SUMMARY OF MODELS FOR REGRESSION

Model	Engine	Dependent Variable (direct labor per engine or equivalent unit)
1	J-79	total hours
2	J-79	fabrication hours
3	J-79	assembly and test hours
4	TF-41	total hours (adjusted for two accounting changes)
5	TF-41	total hours (adjusted for two accounting and one production method changes)
6	F-100	total cost (standard dollars)

Computations for the independent variables, cumulative units and production rate, were similar for all six models.

SUMMARY

Least squares multiple linear regression was used in this research to analyze historical aircraft engine production data. The research hypotheses were tested using the statistical and criterion tests described in this chapter.

The first three research hypotheses were evaluated using statistical tests structured to test the effect of production rate changes on direct labor requirements for aircraft engine production. These tests first evaluated the appropriateness of the full model, then the contribution of the production rate variable. Next, two criterion tests were used, one to test the assumptions necessary to the model, and a second to determine the proportion of variation explained by the regression model. When both the statistical and the criterion tests were passed, the conclusion reached was that the production rate variable was an important explainer of the variation in direct labor requirements for aircraft engine production.

The final hypothesis, that the full model was a better predictor than the reduced model, was evaluated by

simulating predictions of the last 12 data points collected in each data set. The evaluation was a subjective procedure concerned with comparing the predictive abilities of the two models. The full model was judged a better predictor if its predictions deviated from observed values to a lesser degree than the predictions of the reduced model, and if it was subjectively judged to be of practical significance.

In order to draw conclusions with any degree of certainty, the assumptions necessary to the regression model must hold. Also, the methodology contains certain limitations which must be considered throughout the analysis. Due to their importance, a summary of these assumptions and limitations follows.

<u>Assumptions</u>

- Historical data used in the research were accurate.
- No significant loss of data precision was induced by the logarithmic transformation of the data used to facilitate multiple linear regression.
- Data measurements and transformations were accurate.
- The appropriateness of the model was not impaired by multicollinearity between the independent variables.

Limitations

- Subjective analysis was required to assess the validity of the assumptions concerning error terms.
- The utility of the full model as a predictive device must be subjectively determined based on the dollar value of the applicable program versus the increase in predictive ability over the reduced model.
- The ranges of observations in the data bases constrain the utility of the results of this research.

CHAPTER IV

DATA ANALYSIS AND EVALUATION

This chapter provides the results of the methodology described in Chapter III, as applied to the J-79, TF-41, and F-100 production program data sets. It is divided into four sections. The first section reports on the analysis of the three sets of J-79 data with Models 1 through 3[5]. Two different sets of TF-41 data are analyzed with Models 4 and 5 in the second section, while the single F-100 data set is examined with Model 6 in the third. The last section summarizes the analysis by comparing and contrasting the research results of the separate production programs in light of the differences in data and the unique aspects of each program.

THE J-79 PROGRAM

The data obtained on the J-79 production program, unlike those of the other two aircraft engines examined in this research, were provided in separate categories of fabrication data and test and assembly data. This breakout was desired so that differing effects of the production rate on the two different aspects of labor, if present, could be

---

[5]Regression Models 1 through 6 were summarized in Table 4.

detected. Offsetting this attractive attribute of the J-79 data, however, was the data's position near the end of a long production program. The manufacturer's computerized records, from which the data had been extracted, had only been maintained during a period covering unit number 13172 through unit number 14075, a period when production had been slowed as the program approached termination. At the onset of the analysis, then, it was anticipated that the learning curve effect of the cumulative output variable, intuitively prominent when data for an entire production program is available, would be minimal.

Total direct labor data, fabrication direct labor data, and test and assembly direct labor data provided values for use in Models 1, 2, and 3 respectively. The values for the variables in the models were calculated as described in the data collection and treatment section of Chapter III. Both the full and reduced forms of the models were then regressed, and analysis was performed on the results[6]. Specifically, the regression results were

---

[6]The regression analysis in this research was performed through use of a modified version of Smith's FORTRAN IV program (11:147-153). The modified program is listed in Appendix A. A similar program is available for use by government price analysts in the COPPER IMPACT Library under the file name PRODRATE.

evaluated for the three models using the two statistical tests and two criterion tests contained in each model's initial research hypothesis. None of the three models passed all four tests, so the fourth research hypothesis was not evaluated for any of the J-79 data.

Analysis of Research Hypothesis One

To facilitate understanding and provide continuity, the statistical hypotheses and criterion tests for research hypothesis one are restated in summary form. A discussion of test results is then presented.

Research Hypothesis One. The production rate explains a significant portion of the variation in total direct labor requirements for aircraft engine production when included in an appropriate model.

Statistical Hypothesis One(A). $H_0: B_1$ and $B_2 = 0$; $H_1: B_1 \neq 0$ and/or $B_2 \neq 0$. Reject H_0 if the F ratio is greater than F_c.

Statistical Hypothesis One(B). $H_0: B_2 = 0$; $H_1: B_2 \neq 0$. Reject H_0 if the F* ratio is greater than F_c.

Criterion Test One(A). The model cannot be deemed appropriate unless assumptions about constant variance of residuals, residual independence, and normal distribution of

residuals cannot be rejected using appropriate tests.

Criterion Test One(B). The model cannot be deemed appropriate unless the R^2(actual) is greater than 80 percent.

Model 1 Test Results. Test results for research hypothesis one are presented in Table 5 for Model 1. For clarity, the model and its variables are briefly described as follows:

Reduced Model:

$$Y_t = B_0 \cdot X_1^{B_1}$$

or in logarithmic form:

$$\log(Y_t) = \log(B_0) + B_1 \cdot \log(X_1).$$

Full Model:

$$Y_t = B_0 \cdot X_1^{B_1} \cdot X_2^{B_2}$$

or in logarithmic form:

$$\log(Y_t) = \log(B_0)+B_1 \cdot \log(X_1)+B_2 \cdot \log(X_2).$$

where:

Y_t represents unit cost (total direct labor hours/equivalent unit/accounting month),

X_1 represents cumulative output as of the last day of each accounting month, and

X_2 represents daily average production rate/accounting month.

As depicted in Table 5, Model 1 was not validated for further testing under research hypothesis four. Both the full and the reduced forms of the model proved to be

TABLE 5

RESEARCH HYPOTHESIS ONE

RESULTS FOR GENERAL ELECTRIC

MODEL 1

TESTED ITEMS	REDUCED MODEL	FULL MODEL
Estimated B_0	Masked[a]	Masked
Estimated B_1	Masked	Masked
Estimated B_2	Masked	Masked
F Ratio	4.12	3.53
F Critical (2, 52)	--	3.15
Statistical Hypothesis 1A	--	Passed
F*	--	2.78
F Critical (1, 52)	--	4.00
Statistical Hypothesis 1B	--	Failed
Criterion Test 1A	--	Failed
R^2	.076	.126
R^2(Actual)	.085	.134
Criterion Test 1B	--	Failed

[a]Certain table values have been masked in the published version of this thesis, because these data elements are considered proprietary by the manufacturer.

significant explainers of the variation in the dependent variable; however, the production rate variable did not add significantly to this explanatory ability. Furthermore, the assumption concerning independence of residuals was rejected based on the appropriate test, and the R^2 (actual) value of the model was unacceptably low. The rejection of Model 1 as an appropriate model was anticipated as mentioned previously in this section because the data used for the analysis was from a period late in the production program, where the learning curve effect was negligible.

Analysis of Research Hypothesis Two

The statistical hypotheses and criterion tests for research hypothesis two will now be restated in summary form, followed by a discussion of test results.

Research Hypothesis Two. The production rate explains a significant portion of the variation in fabrication direct labor requirements for aircraft engine production when included in an appropriate model.

Statistical Hypothesis Two(A). $H_0: B_1$ and $B_2=0$; $H_1: B_1 \neq 0$ and/or $B_2 \neq 0$. Reject H_0 if the F ratio is greater than F_c.

Statistical Hypothesis Two(B). $H_0: B_2=0$; $H_1: B_2 \neq 0$. Reject

H_0 if the F* ratio is greater than F_c.

Criterion Test Two(A). As with research hypothesis one, the model cannot be deemed appropriate unless the necessary assumptions concerning residuals cannot be rejected.

Criterion Test Two(B). The model cannot be deemed appropriate unless the R^2(actual) is greater than 80 percent.

Model 2 Test Results. Test results for research hypothesis two are presented in Table 6 for Model 2. The model and its variables are briefly described as follows:

Reduced Model:

$$Y_f = B_0 \cdot X_1^{B_1}$$

or in logarithmic form:

$$\log(Y_f) = \log(B_0) + B_1 \cdot \log(X_1).$$

Full Model:

$$Y_f = B_0 \cdot X_1^{B_1} \cdot X_2^{B_2}$$

or in logarithmic form:

$$\log(Y_f) = \log(B_0) + B_1 \cdot \log(X_1) + B_2 \cdot \log(X_2).$$

where:

Y_f represents unit cost (fabrication direct labor hours/equivalent unit/accounting month),

X_1 represents cumulative output as of the last day of each accounting month, and

TABLE 6

RESEARCH HYPOTHESIS TWO

RESULTS FOR GENERAL ELECTRIC

MODEL 2

TESTED ITEMS	REDUCED MODEL	FULL MODEL
Estimated B_0	Masked	Masked
Estimated B_1	Masked	Masked
Estimated B_2	Masked	Masked
F Ratio	3.16	3.24
F Critical (2, 52)	--	3.15
Statistical Hypothesis 2A	--	Passed
F*	--	3.18
F Critical (1, 52)	--	4.00
Statistical Hypothesis 2B	--	Failed
Criterion Test 2A	--	Passed
R^2	.060	.117
R^2(Actual)	.067	.125
Criterion Test 2B	--	Failed

X_2 represents daily average production rate/ accounting month.

Model 2 results, as shown in Table 6, were similar to the Model 1 results for the two statistical tests; that is, the model was a significant explainer of variation, but the production rate variable did not add significantly to the explanatory ability. The assumptions concerning residuals could not be rejected for Model 2. As with Model 1, the R^2(actual) value was far below the established criterion of .80. As a consequence of these results, Model 2 could not be validated for further testing under research hypothesis four. The performance of Model 2 in the regression analysis reflected the close relationship of the Model 2 data to the Model 1 data, in that fabrication direct labor hours were a large percentage of total direct labor hours.

Analysis of Research Hypothesis Three

The statistical hypotheses and criterion tests for research hypothesis three will now be restated in summary form, and test results will be briefly discussed.

Research Hypothesis Three. The production rate explains a significant portion of the variation in assembly and test direct labor requirements for aircraft engine production

when included in an appropriate model.

Statistical Hypothesis Three(A). $H_0: B_1$ and $B_2 = 0$; $H_1: B_1 \neq 0$ and/or $B_2 \neq 0$. Reject H_0 if the F ratio is greater than F_c.

Statistical Hypothesis Three(B). $H_0: B_2 = 0$; $H_1: B_2 \neq 0$. Reject H_0 if the F* ratio is greater than F_c.

Criterion Test Three(A). As with research hypothesis one, the model cannot be deemed appropriate unless the necessary assumptions concerning residuals cannot be rejected.

Criterion Test Three(B). The model cannot be deemed appropriate unless the R^2 (actual) is greater than 80 percent.

Model 3 Test Results. Test results for research hypothesis three are presented in Table 7 for Model 3. The model and its variables are briefly described as follows:

Reduced Model:

$$Y_a = B_0 \cdot X_1^{B_1}$$

or in logarithmic form:

$$\log(Y_a) = \log(B_0) + B_1 \cdot \log(X_1).$$

Full Model:

$$Y_a = B_0 \cdot X_1^{B_1} \cdot X_2^{B_2}$$

or in logarithmic form:

$$\log(Y_a) = \log(B_0) + B_1 \cdot \log(X_1) + B_2 \cdot \log(X_2).$$

TABLE 7

RESEARCH HYPOTHESIS THREE

RESULTS FOR GENERAL ELECTRIC

MODEL 3

TESTED ITEMS	REDUCED MODEL	FULL MODEL
Estimated B_0	Masked	Masked
Estimated B_1	Masked	Masked
Estimated B_2	Masked	Masked
F Ratio	2.21	1.42
F Critical (2, 52)	--	3.15
Statistical Hypothesis 3A	--	Failed
F*	--	0.64
F Critical (1, 52)	--	4.00
Statistical Hypothesis 3B	--	Failed
Criterion Test 3A	--	Passed
R^2	.042	.055
R^2(Actual)	.063	.063
Criterion Test 3B	--	Failed

where:

Y_a represents unit cost (assembly and test direct labor hours/equivalent unit/accounting month),

X_1 represents cumulative output as of the last day of each accounting month, and

X_2 represents daily average production/accounting month.

As depicted in Table 7, Model 3 was not validated for further testing under research hypothesis four. Neither the full nor the reduced form of the model proved to be a significant explainer of the variation in the dependent variable. Two direct consequences of this result were that (1) the addition of the production rate variable could not, and in fact did not, add significantly to the explanatory ability of the model; and, (2) the R^2(actual) value was unacceptably low. The model passed the tests concerning the assumptions about residuals.

THE TF-41 PROGRAM

The TF-41 program data used in this research covered the entire 11-year production history of the engine to date, beginning with unit number one in 1968. With the entire production history reflected in the data, the full

learning curve effect of the cumulative output variable was shown. A breakout between fabrication direct labor and assembly and test direct labor was not available, so analysis of the explanatory ability of the appropriate models was confined to research hypothesis one. Two different models were used in this analysis, each reflecting different transformations of the data as discussed in Chapter III.

Analysis of Research Hypothesis One

To provide full understanding and continuity, the statistical hypotheses and criterion tests for research hypothesis one are again summarized prior to a discussion of test results for Model 4 and Model 5.

Research Hypothesis One. The production rate explains a significant portion of the variation in total direct labor requirements for aircraft engine production when included in an appropriate model.

Statistical Hypothesis One(A). $H_0: B_1$ and $B_2 = 0$; $H_1: B_1 \neq 0$ and/or $B_2 \neq 0$. Reject H_0 if the F ratio is greater than F_c.

Statistical Hypothesis One(B). $H_0: B_2 = 0$; $H_1: B_2 \neq 0$. Reject H_0 if the F* ratio is greater than F_c.

Criterion Test One(A). The model cannot be deemed appropriate

unless assumptions about constant variance of residuals, residual independence, and normal distribution of residuals cannot be rejected using appropriate tests.

Criterion Test One(B). The model cannot be deemed appropriate unless the R^2(actual) is greater than 80 percent.

Model 4 Test Results. Test results for research hypothesis one are presented in Table 8 for Model 4. To insure understanding, the model and its variables are briefly described as follows:

Reduced Model:

$$Y_t = B_0 \cdot X_1^{B_1}$$

or in logarithmic form:

$$\log(Y_t) = \log(B_0) + B_1 \cdot \log(X_1).$$

Full Model:

$$Y_t = B_0 \cdot X_1^{B_1} \cdot X_2^{B_2}$$

or in logarithmic form:

$$\log(Y_t) = \log(B_0) + B_1 \cdot \log(X_1) + B_2 \cdot \log(X_2).$$

where:

Y_t represents unit cost (total direct labor hours/ equivalent unit/accounting month, adjusted for two accounting changes[7]),

---

[7]As described in the data collection and treatment section of Chapter III.

TABLE 8

RESEARCH HYPOTHESIS ONE

RESULTS FOR ALLISON

MODEL 4

TESTED ITEMS	REDUCED MODEL	FULL MODEL
Estimated B_0	Masked	Masked
Estimated B_1	Masked	Masked
Estimated B_2	Masked	Masked
F Ratio	141.79	169.65
F Critical (2, 126)	--	3.07
Statistical Hypothesis 1A	--	Passed
F*	--	92.68
F Critical (1, 126)	--	3.92
Statistical Hypothesis 1B	--	Passed
Criterion Test 1A	--	Failed
R^2	.533	.734
R^2(Actual)	.702	.832
Criterion Test 1B	--	Passed

X_1 represents cumulative output as of the last day of each accounting month, and

X_2 represents daily average production rate/ accounting month.

The results depicted in Table 8 show that Model 4 could not be validated for further testing under research hypothesis four. Both the full and the reduced forms of the model proved to be significant explainers of the variation in the dependent variable, and the addition of the production rate variable added significantly to this explanatory ability. Additionally, the R^2(actual) value of .832 for the full model was above the criterion limit. However, Model 4 failed to pass the tests concerning assumptions of independence and normal distribution of residuals. The plot of residual values against the cumulative output variable revealed a strong pattern, as depicted in Figure 1. In this pattern the residual values decreased as cumulative production increased to the 100th data point, then jumped drastically to high positive values. The jump coincided with the production method change which occurred beginning with data point 101.

Model 5 Test Results. Test results for research hypothesis one are presented in Table 9 for Model 5. For clarification,

FIGURE 1

PLOT OF RESIDUALS FOR MODEL 4

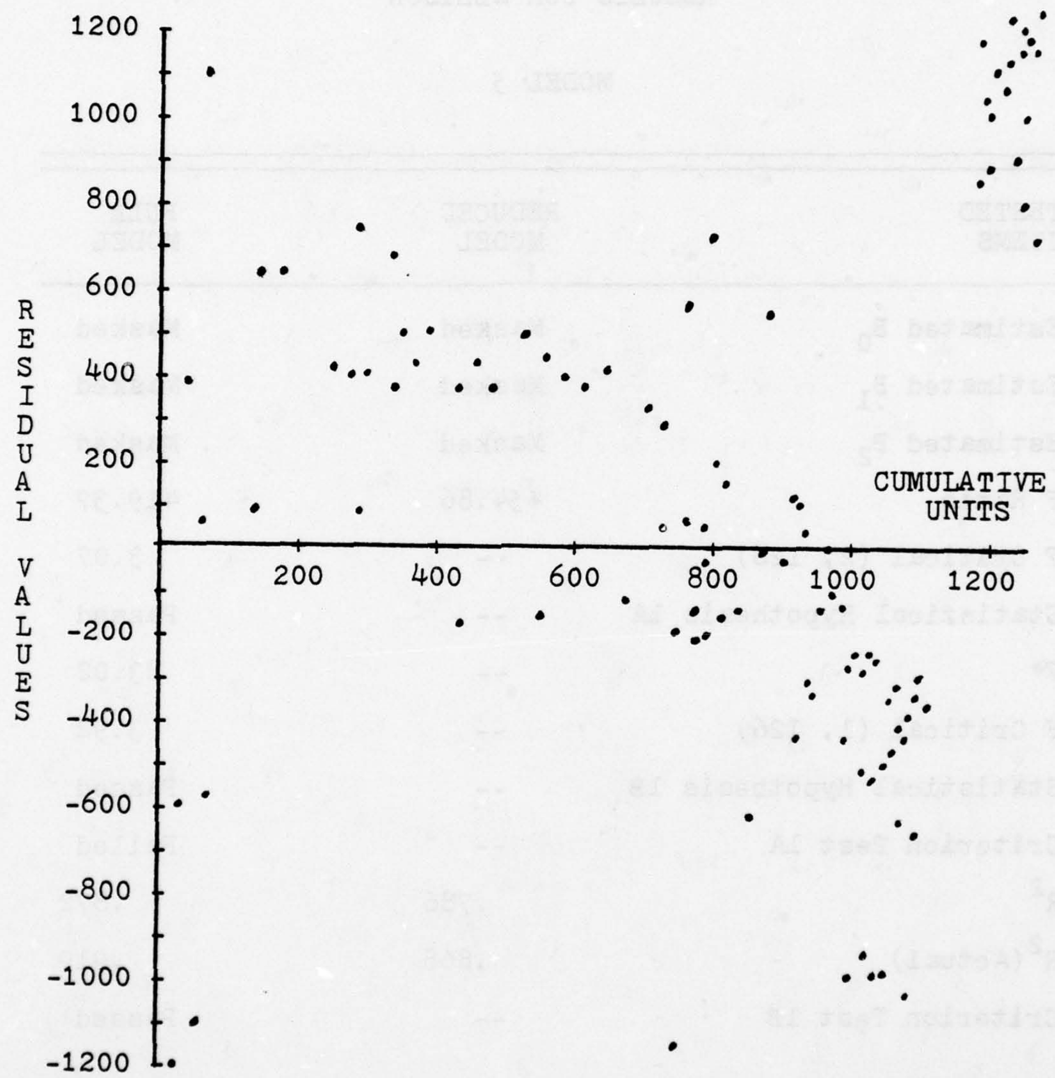


TABLE 9

RESEARCH HYPOTHESIS ONE

RESULTS FOR ALLISON

MODEL 5

TESTED ITEMS	REDUCED MODEL	FULL MODEL
Estimated B_0	Masked	Masked
Estimated B_1	Masked	Masked
Estimated B_2	Masked	Masked
F Ratio	454.86	419.37
F Critical (2, 126)	--	3.07
Statistical Hypothesis 1A	--	Passed
F*	--	83.02
F Critical (1, 126)	--	3.92
Statistical Hypothesis 1B	--	Passed
Criterion Test 1A	--	Failed
R^2	.786	.872
R^2 (Actual)	.868	.919
Criterion Test 1B	--	Passed

the model and its variables are briefly described as follows:

Reduced Model:

$$Y_t = B_0 \cdot X_1^{B_1}$$

or in logarithmic form:

$$\log(Y_t) = \log(B_0) + B_1 \cdot \log(X_1).$$

Full Model:

$$Y_t = B_0 \cdot X_1^{B_1} \cdot X_2^{B_2}$$

or in logarithmic form:

$$\log(Y_t) = \log(B_0) + B_1 \cdot \log(X_1) + B_2 \cdot \log(X_2).$$

where:

Y_t represents unit cost (total direct labor hours/ equivalent unit/accounting month, adjusted for two accounting and one production method changes[8]),

X_1 represents cumulative output as of the last day of each accounting month, and

X_2 represents daily average production rate/accounting month.

Model 5 results, as depicted in Table 9, show that the model passed both statistical tests and one criterion test but failed the criterion test concerning assumptions about residuals. Consequently, it was not validated for

---

[8]As described in the data collection and treatment section of Chapter III.

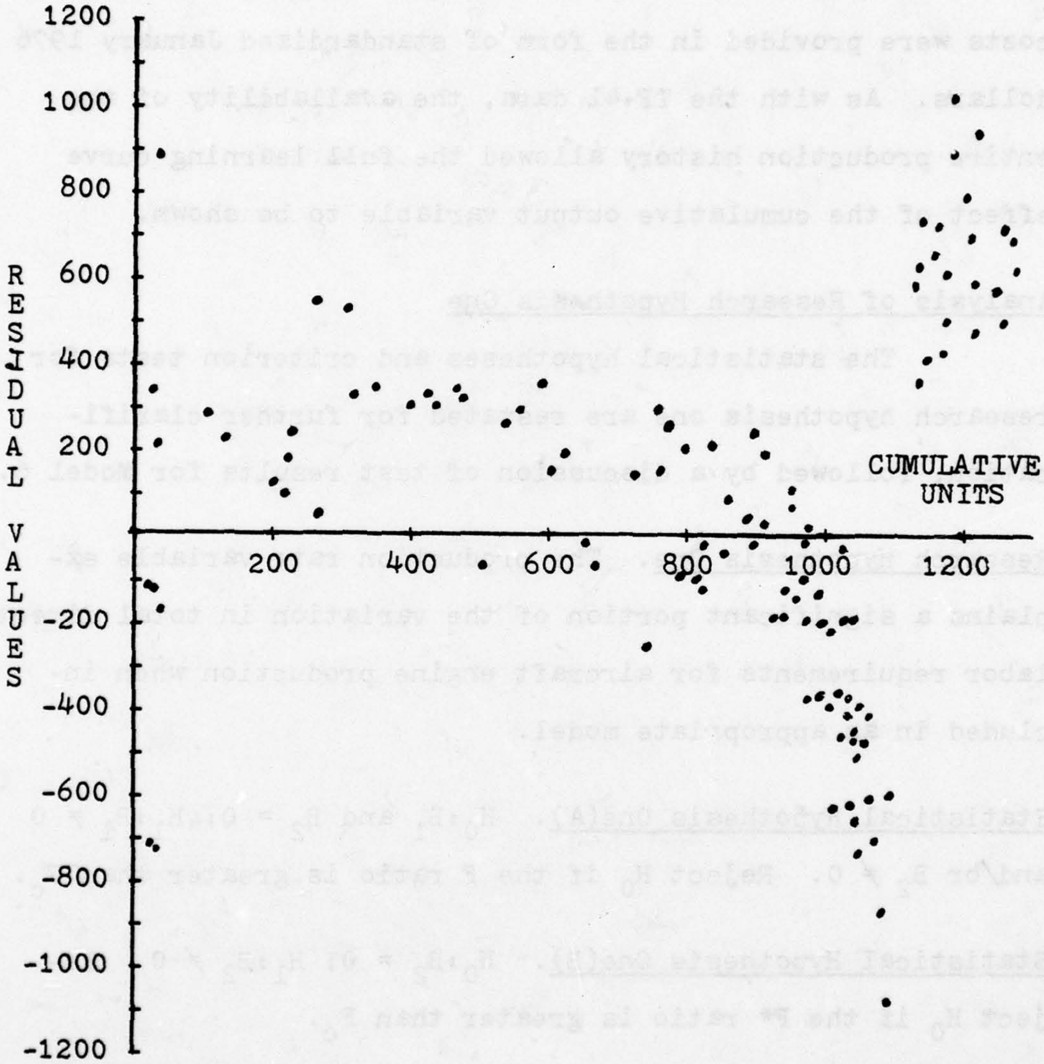
further testing under research hypothesis four. Model 5's failure to pass criterion test one (A) was due to the rejection of the assumption of independence of residuals. As with Model 4, the plot of residual values against the cumulative output variable revealed a pattern which coincided with the production method change at data point 101. This pattern is shown in Figure 2. The pattern was less pronounced than with Model 4, because data used in Model 5 were transformed to compensate for the production method change. The fact that the pattern among residuals persisted highlighted a limitation of linear regression analysis as performed in this research. Specifically, breaks in production, such as resulted from the production method change in the TF-41 fabrication process, cannot be accommodated by either the reduced model or the full model. Techniques using a zero-one variable in the regression equation can adequately adjust the equation to compensate for the effect of a production break. Application of such techniques is beyond the scope of this research.

THE F-100 PROGRAM

The data for the F-100 program which was used in this research covered the entire production history of the engine to date, where 1090 units were produced from May 1973

FIGURE 2

PLOT OF RESIDUALS FOR MODEL 5



to January 1979. No breakout between fabrication direct labor and assembly and test direct labor was available, so analysis of the explanatory ability of the appropriate models was limited to research hypothesis one. All direct labor costs were provided in the form of standardized January 1976 dollars. As with the TF-41 data, the availability of the entire production history allowed the full learning curve effect of the cumulative output variable to be shown.

Analysis of Research Hypothesis One

The statistical hypotheses and criterion tests for research hypothesis one are restated for further clarification, followed by a discussion of test results for Model 6.

Research Hypothesis One. The production rate variable explains a significant portion of the variation in total direct labor requirements for aircraft engine production when included in an appropriate model.

Statistical Hypothesis One(A). $H_0: B_1$ and $B_2 = 0$; $H_1: B_1 \neq 0$ and/or $B_2 \neq 0$. Reject H_0 if the F ratio is greater than F_c.

Statistical Hypothesis One(B). $H_0: B_2 = 0$; $H_1: B_2 \neq 0$. Reject H_0 if the F* ratio is greater than F_c.

Criterion Test One(A). The model cannot be deemed appropriate

unless the assumptions concerning residuals are met.

Criterion Test One(B). The model cannot be deemed appropriate unless the R^2(actual) is greater than 80 percent.

Model 6 Test Results. Test results for research hypothesis one are presented in Table 10 for Model 6. The model and its variables are briefly described as follows:

Reduced Model:

$$Y_t = B_0 \cdot X_1^{B_1}$$

or in logarithmic form:

$$\log(Y_t) = \log(B_0) + B_1 \cdot \log(X_1).$$

Full Model:

$$Y_t = B_0 \cdot X_1^{B_1} \cdot X_2^{B_2}$$

or in logarithmic form:

$$\log(Y_t) = \log(B_0) + B_1 \cdot \log(X_1) + B_2 \cdot \log(X_2).$$

where:

Y_t represents unit cost (total direct labor dollars/equivalent unit/accounting month),

X_1 represents cumulative output as of the last day of each accounting month, and

X_2 represents daily average production rate/month.

The results depicted in Table 10 show that Model 6 passed both statistical tests and both criterion tests, and was validated for further testing under research hypothesis

TABLE 10

RESEARCH HYPOTHESIS ONE

RESULTS FOR PRATT AND WHITNEY

MODEL 6

TESTED ITEMS	REDUCED MODEL	FULL MODEL
Estimated B_0	Masked	Masked
Estimated B_1	Masked	Masked
Estimated B_2	Masked	Masked
F Ratio	215.51	484.66
F Critical (2, 64)	--	3.15
Statistical Hypothesis 1A	--	Passed
F*	--	169.18
F Critical (1,64)	--	4.00
Statistical Hypothesis 1B	--	Passed
Criterion Test 1A	--	Passed
R^2	.777	.941
R^2 (Actual)	.628	.902
Criterion Test 1B	--	Passed

four. F ratio tests for statistical hypotheses one(A) and one(B) were significant to a level of confidence beyond .001. Also Model 6 passed all tests on the assumptions concerning the residual values. The R^2(actual) value of .90 was well above the selected criterion of .80 under criterion test one(B).

Analysis of Research Hypothesis Four

The computer program listed in Appendix A was used to perform the predictive ability test on Model 6. The researcher can select an option in the program which performs stepwise truncation of individual data points, beginning with the last point in the data set and continuing in reverse sequence. This option simulates prediction of direct labor requirements for each truncated data point, then compares the predicted value and computes the resultant difference between the values, both as an absolute deviation and as a percent deviation from the observed value.

In this research, the aforementioned option was exercised for the Model 6 data, truncating the last 12 points in the data set. A summary of program results is depicted in Table 11, showing the percent deviation of each predicted value from the corresponding observed value,

TABLE 11

SHORT RANGE PREDICTIVE ABILITY COMPARISON

FOR PRATT AND WHITNEY MODEL 6[a]

Data Point	Reduced Model	Full Model
	Percent Deviation	Percent Deviation
64	9.67	-1.63
63	7.01	-5.28
62	-2.07	0.24
61	19.36	-2.42
60	7.29	-1.59
59	8.17	-2.26
58	9.07	1.07
57	14.31	2.14
56	11.63	-3.17
55	9.11	5.59
54	8.68	-1.06
53	10.63	11.44

[a]The first 52 of the 64 values in the data set were used in the regression to predict the last 12 values.

for each of the 12 truncated data points, for both the full and reduced models. In this manner the prediction of direct labor requirements for the most recent 12 months' production was simulated, using the production history of the engine with those final 12 data points deleted. It must be noted that, to compute the aggregate (12-month) deviation values for the full and reduced models, the monthly deviations shown in Table 11 were weighted in proportion to the number of units completed during each month.

Research Hypothesis Four. For twelve months into the future, the predictive ability of the full model is better than the predictive ability of the reduced model.

Statistical Hypothesis Four. $H_0: |\bar{D}_R| \leq |\bar{D}_F|$; $H_1: |\bar{D}_R| > |\bar{D}_F|$. Reject H_0 if $t > t_c$ (.05).

Criterion Test Four. The full model is considered of practical value if its 12-month aggregate deviation is subjectively judged to be (1) sufficiently smaller than that of the reduced model, and (2) sufficiently small in an absolute sense.

Model 6 Test Results. As depicted in Table 12, Model 6 passed both the statistical test and the criterion test, and was consequently judged to be a better predictor than

TABLE 12

RESEARCH HYPOTHESIS FOUR

RESULTS FOR PRATT AND WHITNEY

MODEL 6

TESTED ITEMS	REDUCED MODEL	FULL MODEL
Average Absolute Deviation	5144.81	3252.98
Variance	7,849,467.56	6,687,281.06
t Test Statistic	---	8.93
t Critical	---	1.80
Statistical Hypothesis Four Result		Reject H_0
Percent Deviation from Actual 12-Month Direct Labor Figure	8.48	.32
Criterion Test Four Result		Passed

the reduced model. The t-test results were significant beyond the .001 level of confidence, yet the most impressive performance of the full model was in its evaluation using the criterion test. In this test the full model's accuracy in predicting the direct labor costs for the most recent 12 months of production was found to be within one-third of one percent of the actual value, while the reduced model's prediction deviated from the actual by more than eight percent. The inescapable conclusion was that the full model was not only a better predictor than the reduced model, but was of strong practical value.

COMPARATIVE ANALYSIS AND SUMMARY

The regression analysis with the three sets of engine production data produced three distinctly different results. A comparative analysis of the J-79, TF-41, and F-100 engine data and production programs will investigate these differences. A summary and evaluation of the applicability of the production rate model to the three sets of data will then complete Chapter IV.

Comparative Analysis

A thorough analysis of the results of the hypothesis testing described in this chapter revealed two significant

AD-A077 649 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL--ETC F/G 5/1
AN INVESTIGATION OF CHANGES IN DIRECT LABOR REQUIREMENTS RESULT--ETC(U
SEP 79 M W CROZIER, E J MCGANN
UNCLASSIFIED AFIT-LSSR-22-79B NL

2 OF 2
AD
A077649

END
DATE
FILMED
1 — 80
DDC

findings concerning applicability of the models tested. The first was that the tremendous complexity of modern aircraft engine production programs limited the applicability of the learning curve in either its reduced or full form. The second was that a substantial portion of the production history must be included in a data set in order to effectively measure the applicability of the appropriate models. These two findings will now be discussed.

Program Complexity. Modern aircraft propulsion systems are composed of thousands of parts, many of which are precision tooled, complex in nature, and fabricated from advanced alloys. Many parts are common to several programs; others are unique. Additionally, a mature production program often produces several engine models, each differing in varying degrees from the others. The extensive nature of the fabrication and assembly tasks involved in the process affords management constant opportunity to alter make-or-buy policies and in-plant production methods. This dynamic state of production places in potential jeopardy the assumptions critical to the successful application of the learning curve which were stated in Chapter I.

The TF-41 program underwent numerous changes

throughout its production history, including the two accounting changes and the major production method change described in the data collection and treatment section of Chapter III. The periodic minor changes to improve production methods were anticipated by the model and were an important part of the learning process; however, the major production method change had an adverse impact on the applicability of the model, and in fact caused the failure of both Model 4 and Model 5 under research hypothesis one. The data transformation performed in the Model 5 analysis to account for the production method change did improve the results in comparison to Model 4's, but was not sufficient to allow the model to pass the test of the assumption of independence of residuals. The production method change had introduced an anomaly in direct labor costs which the full and reduced models of the learning curve could not adequately explain. This anomaly caused the residual values to appear to be dependent upon the cumulative units of production, which in turn caused Model 5's rejection under research hypothesis one. Yet, even if the model had somehow been validated under research hypothesis one, it would still have failed the criterion test under research hypothesis four due to a full 14 percent deviation from the 12-month aggregate actual direct labor costs. This large deviation was

another effect of the production method change.

The other two data sets did not reflect the same impact of program complexity, for differing reasons. The F-100 program was relatively new with no major program changes. The J-79 data set was at the end of a lengthy production run, where any major program changes had already occurred.

Production History. The discussion on program complexity in the preceding section reflected the hazards of utilizing program data from the most dynamic, substantial portion of the production program. However, this research also highlights the importance of obtaining data from a major portion of the production program in order to capture the full learning curve effect.

Problems encountered in analyzing the applicability of the appropriate learning curve models to a small segment of production data were shown in the analysis of the J-79 program. The only J-79 data available for analysis in this research was a five-year segment near the end of the production program, which included the last 900 units of an approximate 14,000 unit program. Viewed graphically, this was a nearly horizontal segment of the entire production history learning curve. As a result, the amount of variation in

direct labor costs which was explained by the learning curve effect was not a significant improvement over the amount explained by the simple arithmetic average. Of interest was that for the J-79 program, the production rate independent variable alone was a statistically significant explainer of variation in total direct labor costs. However, considering the negligible effect of the cumulative output independent variable, the overall explanatory ability of the model was unacceptably low. In contrast, the F-100 and TF-41 data encompassed the entire histories of the programs, and in both cases the effect of the learning curve was prominent.

Summary

The preceding analyses of the three aircraft engine production programs were based on the methodology and data treatment as explained in Chapter III. Research hypothesis one was only accepted for Model 6 using the F-100 data set, and research hypotheses two and three were not accepted for the models tested. The production rate variable was an important explainer of variation in direct labor costs for three of the six models evaluated. The increase in R^2(actual) was greater than 5 percent in five of the models as shown in Table 13. The failure of the J-79 and TF-41 program

TABLE 13

INCREASES IN R^2(ACTUAL) FOR ALL SIX MODELS

TESTED AFTER INCLUSION OF PRODUCTION RATE

MODEL	R^2 REDUCED	R^2 FULL	ΔR^2
1	.085	.134	.051
2	.067	.125	.058
3	.063	.063	.000
4	.702	.832	.130
5	.868	.919	.051
6	.628	.902	.274

models to be accepted under any of the first three research hypotheses was due to program complexity and insufficient production history respectively, as discussed in the preceding section of this chapter.

Model 6, developed for the F-100 data set, was also accepted under research hypothesis four. Not only was the predictive ability of the full model statistically greater than that of the reduced model at the .05 level of significance, but the superiority of the full model was also clearly demonstrated by the subjective criterion tests.

The analysis of the J-79, TF-41, and F-100 engine production programs performed in this research thus met the research objectives of determining the effect of production rate changes on production direct labor requirements and evaluating the predictive ability of a cumulative production and production rate model. Of the three engines, only the F-100 engine passed all the tests.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Inflation, energy shortages and other current economic forces have more sharply focused the attention of the American public on the manner in which tax dollars are spent. This public pressure for reduced government spending has been directed extensively toward major system acquisitions by the Department of Defense, in part due to the high visibility of these sytems caused by the Congressional budgetary process.

The reduction of the DOD budget in real terms during a time of greater complexity and expense of major defense systems, combined with the strong public scrutiny as to how the budget is spent, points to the critical need for improved cost estimating techniques and procedures. Direct labor costs drive a substantial portion of total major system acquisition costs, making direct labor a key cost element. The learning curve is often employed to estimate direct labor costs.

LITERATURE REVIEW

Learning curve models have served as techniques for estimating direct labor costs since their development in

the aerospace industry of the World War II era. Following the introduction of the basic learning curve model, several variations were developed in a continuing quest for greater accuracy in cost forecasting. The most promising of these variations has been the production rate model.

Most researchers who studied the effect of production rate changes concluded that the production rate was a significant explainer of direct labor costs. The most conclusive of these studies was that performed by Smith, who developed a generally accepted technique for estimating airframe direct labor requirements which systematically employed a production rate variable in the learning curve model. Smith's production rate model focused on airframes; this research extended the investigation of production rate from airframes to aircraft engines.

THE MODEL

Smith's production rate model, replicated by Congleton and Kinton, was a modification of the model developed by Orsini. The full model used in this research, an adaptation of Smith's model, is as follows:

$$Y = B_0 \cdot X_1^{B_1} \cdot X_2^{B_2} \cdot 10^e$$

where the variables are described in general terms as follows:

Y represents the unit average direct labor hours (or proxy),

X_1 represents the cumulative output of engines or equivalent units,

X_2 represents the production rate,

e represents the variation which is left unexplained by the variables in the model, and

B_0, B_1, and B_2 are regression coefficients.

The logarithmic form of the full model is:

$$\log Y = \log B_0 + B_1 \log X_1 + B_2 \log X_2 + e.$$

The reduced model used in this research is the basic learning curve model, identical to the full model both before and after transformation to logarithmic form, except the reduced model does not include the production rate variable.

RESEARCH OBJECTIVES

The two primary objectives of this research were (1) to determine if changes in the rate of production affect total direct labor requirements per engine, and (2) to determine if the production rate model was better than the basic learning curve model in predicting direct labor requirements for aircraft engine production.

METHODOLOGY

The methodology designed to meet the research objectives was based on linear regression of the logarithmic forms of the full and reduced models. Data for use in the regression analysis were obtained from three aircraft engine production programs. Each data set contained unique aspects requiring individual treatment as described in Chapter III, and specific models for regression were developed for each data set. Models were analyzed to determine first if the production rate was an important explainer of direct labor hours, and if so, whether the full model was a better predictor of direct labor hours than the reduced model. The hypotheses which comprised these steps are described next.

Research hypothesis one stated that production rate was an important explainer of total direct labor hours when included in an appropriate model. Two statistical tests were employed. First, the F-test was used to determine if a relationship existed between the dependent variable and the independent variables, cumulative output and production rate. Next, the F-test was again used to determine if the production rate was significant in explaining additional variation in the dependent variable.

Following the two statistical tests, two criterion

tests were performed. The first criterion test evaluated the assumptions of constant variance of residuals, residual independence, and normal distribution of residuals, three assumptions which were necessary for the model to be deemed appropriate. The second criterion test required R^2(actual) to exceed 80 percent, an arbitrary measure of the model's ability to explain a sufficient amount of variation in the dependent variable. Models which passed both statistical and both criterion tests were validated for further testing under research hypothesis four.

Research hypotheses two and three were identical to research hypothesis one, except that direct fabrication labor requirements and assembly and test labor requirements respectively were substituted for total direct labor requirements. As in research hypothesis one, models were required to pass both statistical and both criterion tests in order to be validated for further testing under research hypothesis four.

Tests under research hypothesis four were used to measure and compare the predictive abilities of the full and reduced forms of a model which passed the tests of explanatory ability just described. A statistical test and a criterion test were used in this procedure, whereby

the last 12 months, or data points, in the data set were truncated, and the remaining data points were used in a simulation to predict these 12 points. In the statistical test the t-test was used to determine if the average absolute deviation of the full model was significantly less than that of the reduced model. Then the criterion test compared the aggregate 12-month deviation of the two models in a test of practical significance.

With the exception of the tests performed on assumptions of residuals and the subjective evaluation of practical importance in criterion test four, all test results were obtained through use of the computer program PRODRATE listed in Appendix A.

CONCLUSIONS

Two primary conclusions were reached in the accomplishment of the objectives of this research. Production rate was found to be a significant explainer of variation in direct labor hours in three of the six cases examined. Thus, the production rate variable may or may not be a significant explainer, depending upon the particular program. In the one case where the test of predictive ability was subsequently performed, the full model was found to be

a much better predictor than the reduced model. This result demonstrates that the full model has potential to be a much better predictor than the reduced model, again depending upon the particular program.

Several additional conclusions were reached, each of which will now be discussed. The first additional conclusion is that the complexity of aircraft engine production programs limits the applicability of use of either the full or reduced models used in this research. The large number of parts in an aircraft engine affords management constant opportunity to alter make-or-buy policies and in-plant production methods. Additionally, model changes can frequently occur. These policy and model changes can impact the production history in such a way that neither the full nor reduced model is an appropriate best explainer of variation.

In this research, problems were encountered in analyzing the applicability of the appropriate learning curve models when using a relatively small segment of total program production data. The second additional conclusion is that the engine production data must include a substantial percentage increase in cumulative units of production in order to capture the full effect of the learning curve. Only in this way can the full and reduced models be found appropriate to the data.

Coefficients resulting from each regression analysis performed were unique to that particular data set. These results can be generalized into the obvious conclusion that the regression coefficients are unique to the program for which they were derived. Thus the third additional conclusion is that it is not possible to develop a general model for use in all aircraft engine production programs.

RECOMMENDATIONS

In view of the results obtained for the F-100 engine production program, the most obvious recommendation is that the production rate model be applied in actual direct labor cost estimating for this program. A second and related recommendation is that the production rate model be evaluated for possible use in other ongoing engine production programs. A third recommendation is to extend the scope of this research outward to other major production programs, to investigate the applicability of the production rate model to production of missiles and other major systems.

An additional area for extensive research which was mentioned briefly in this analysis is the adaptation of the model to accomodate changes in production methods and/or policy. Progress in this area would greatly extend the applicability of the model.

In final conclusion, the production rate model as developed by Smith has tremendous potential as a tool for cost estimating in selected aircraft engine production programs.

APPENDIX A

THE COMPUTER PROGRAM PRODRATE

This section lists the computer program PRODRATE. The original program, developed by Lt. Col. Larry L. Smith, was modified by Capt. David Stevens. This modified version, the cumulative production and production rate cost model, was further modified for this research in order to accommodate the aircraft engine data. The actual program used in this research is the program presented in this section.

```
C**********************************************************************************************
C
C     PPPPP  RRRRR  00000  DDD    RRRRR  AAAAA  TTTTT  EEEEE
C     P   P  R   R  O   O  D  D   R   R  A   A    T    E
C     PPPPP  RRRRR  0   0  D   D  RRRRR  AAAAA    T    EEE
C     P      R  R   0   0  D  D   R  R   A   A    T    E
C     P      R   R  00000  DDD    R   R  A   A    T    EEEEE
C
C**********************************************************************************************
C
C                     THE CUMULATIVE PRODUCTION AND PRODUCTION RATE COST MODEL
C
C     THE ORIGINAL PROGRAMMER IS LT COL LARRY L. SMITH [AFIT/LSCM AV# 785-5096] - JAN 1978
C     THIS MODIFIED VERSION WAS PROGRAMMED BY CAPT DAVID Y. STEVENS [ESD/PKG AV# 478-3462] -JUNE 1979
C**********************************************************************************************
 5 FORMAT(1H1,//,100("*"),//,1X,40X,"PRODRATE INSTRUCTIONS",//,1X,100("*"),//,
&  "      THIS PROGRAM IS DESIGNED TO EVALUATE THE VARIATION IN DIRECT LABOR REQUIREMENTS AS A ",/,
&  " FUNCTION OF CUMULATIVE PRODUCTION AND PRODUCTION RATE.  IN ADDITION, THE ANALYST MAY ",/,
&  " COMPARE THE RESULTS OBTAINED FROM THE STANDARD LEARNING CURVE WITH THE RESULTS OBTAINED ",/,
&  " FROM THE CUMULATIVE PRODUCTION AND PRODUCTION RATE MODEL.  THE COST MODELS USED IN THIS ",/,
&  " PROGRAM ARE:",//,
&  "      1.  REDUCED MODEL [STANDARD LEARNING CURVE MODEL]",//,
&  "                Y = B0 + (X1 ** B1) + (10 ** E)",//,
&  "      2.  FULL MODEL [CUMULATIVE PRODUCTION AND PRODUCTION RATE MODEL]",//,
&  "                Y = B0 + (X1 ** B1) + (X2 ** B2) + (10 ** E)",//,
&  "          WHERE: Y                IS THE DIRECT LABOR REQUIREMENTS",/,
&  "                 X1               IS THE CUMULATIVE PRODUCTION PLOT POINT",/,
&  "                 X2               IS THE PRODUCTION RATE PROXY(E.G. EQUIVALENT UNITS PER MONTH)",/,
&  "                 E                REPRESENTS THE ERROR TERM",/,
&  "                 B0, B1, AND B2  ARE  PARAMETERS DETERMINED BY REGRESSION",//,
&  "      DATA ARE INPUT BY READING FROM ANY PROPERLY FORMATTED DATA FILE.  YOUR DATA FILE SHOULD ",/,
&  " BE SAVED TO ANY PERMANENT FILENAME.  YOU WILL BE ASK TO INPUT THE NAME OF YOUR DATA FILE ",/,
&  " AT THE APPROPRIATE STEP IN THE PROGRAM.  THE NAME OF YOUR DATA FILE CAN NOT EXCEED 8 ",/,
&  " CHARACTERS.  THE FIRST LINE OF THE DATA FILE MUST CONTAIN A LINE NUMBER AND THE NUMBER OF ",/,
&  " CASES TO BE READ.  THE DATA IS THEN ENTERED ONE CASE PER LINE IN THE FOLLOWING ORDER: ",/,
&  " LINE NUMBER, OBSERVED DIRECT LABOR REQUIREMENT (Y), CUMULATIVE PRODUCTION PLOT POINT (X1),",/,
&  " AND PRODUCTION RATE PROXY (X2).  THE PROGRAM USES A FREE FIELD READ FORMAT; THEREFORE,",/,
&  " EACH VARIABLE MUST BE SEPARATED BY AT LEAST ONE SPACE (OR OTHER DELIMITER) BUT NO OTHER",/,
&  " SPECIAL FORMAT IS REQUIRED.  AN EXAMPLE OF A DATA FILE WITH 5 CASES IS PRESENTED BELOW:",//,
&  "      100     5",/,
&  "      101     100      9.5       9.5",/,
&  "      102      90     30        20.5",/,
&  "      103      80     55        25",/,
&  "      104      75     82        27",/,
&  "      105      71    113        31",//,
&  " ONE ADVANTAGE OF THIS PROGRAM IS THAT THE RESULTS OBTAINED WILL BE IN THE SAME UNITS AND",/,
&  " FORM AS THE INPUT DATA.  FOR EXAMPLE, IF YOU ARE WORKING IN DIRECT LABOR HOURS PER MONTH",/,
&  " AND EQUIVALENT UNITS, THE RESULTS WILL BE IN TERMS OF THESE UNITS.  ALSO, IF YOU WISH TO USE",/,
&  " A CUMULATIVE AVERAGE APPROACH, ALL YOU NEED DO IS AGGREGATE THE DATA BASE IN THAT MANNER.",//,
&  "      THE PROGRAM BEGINS BY TRANSFORMING THE INPUT DATA TO COMMON LOGARITHMS.  LOG LINEAR",/,
&  " REGRESSION IS THEN PERFORMED AS FOLLOWS:  Y  REGRESSED ON  X1, Y  REGRESSED ON  X2, AND",/,
&  " FINALLY Y  REGRESSED ON BOTH X1 AND X2.  OBSERVED DIRECT LABOR REQUIREMENTS, PREDICTED",/,
&  " DIRECT LABOR REQUIREMENTS, AND RESIDUALS ARE PRINTED IN ORIGINAL (UNTRANSFORMED) FORM FOR",/,
&  " EACH REGRESSION SITUATION.  IN ADDITION, SUMMARY STATISTICS ARE PRINTED FOR EACH MODEL.  THE",/,
&  " SUMMARY STATISTICS INCLUDE TWO COEFFICIENTS OF DETERMINATION R SQUARED LOG AND R SQUARED",/,
&  " ACTUAL.  THE R SQUARED LOG REPRESENTS THE GOODNESS OF FIT OF THE MODEL TO THE TRANSFORMED",/,
&  " DATA (LOG FORM).  THE R SQUARED ACTUAL, ON THE OTHER HAND, IS COMPUTED USING THE",/,
&  " UNTRANSFORMED RESIDUALS, AND IS REPRESENTATIVE OF HOW WELL THE MODEL FITS THE UNTRANSFORMED",/,
&  " DATA.")
```

```
1590 6 FORMAT(1H1,1X,//,
1600  "       SEVERAL OPTIONS ARE AVAILABLE WITHIN THIS PROGRAM AND CAN BE SELECTED BY APPROPRIATE",/,
1610  " ANSWERS TO THE FOLLOWING QUESTIONS:",//,
1620  "       1.  DO YOU WANT TO CHECK DATA AS IT IS READ FROM FILE ....... AND CONVERTED TO",/,
1630  "       LOGARITHMS?",//,
1640  "       YES - WILL CAUSE THE PRINTING OF A LISTING OF THE RATIONAL INPUT DATA AND THE",/,
1650  "             ASSOCIATED LOGARITHMIC VALUES.",//,
1660  "       NO  - SUPPRESSES THIS OPTION.",//,
1670  "       2.  DO YOU WANT A COMPARISON OF THE SHORTRANGE PREDICTIVE ABILITY OF THE TWO MODELS?",//,
1680  "       YES - WILL CAUSE THE PREDICTIVE ABILITY TEST OPTION TO BE ACTIVATED AND THE USER WILL",/,
1690  "             BE ASKED: 'HOW MANY CASES DO YOU WISH TO TRUNCATE?'  THE RESPONSE TO THIS",/,
1700  "             QUESTION MAY BE ANY INTEGER VALUE GREATER THAN OR EQUAL TO 2.  THE PREDICTIVE",/,
1710  "             ABILITY TEST SIMULATES FUTURE PREDICTIONS BY PERFORMING A STEPWISE TRUNCATION OF",/,
1720  "             THE HISTORICAL DATA.  FOR THIS REASON, AN UPPER LIMITATION ON THE NUMBER OF",/,
1730  "             CASES TRUNCATED WOULD BE: ((TOTAL NUMBER OF CASES IN DATA FILE) / 2) - 2",/,
1740  "             FOR EXAMPLE, IF YOUR DATA FILE CONTAINS 50 CASES, YOUR UPPER LIMIT WOULD BE",/,
1750  "             23 CASES.  THIS, OF COURSE, REPRESENTS ONLY THE MAXIMUM NUMBER OF CASES THAT",/,
1760  "             COULD BE TRUNCATED.  IN PRACTICE YOU MAY WANT TO TRUNCATE ONLY A SMALL NUMBER OF",/,
1770  "             CASES.  THUS, IF YOUR DATA IS COLLECTED IN MONTHLY INTERVALS, YOU CAN LOOK AT",/,
1780  "             THE PREDICTIVE ABILITY OF THE FULL AND REDUCED MODELS FOR AN 18 MONTH TIME SPAN BY",/,
1790  "             SPECIFYING '18'.  LIKEWISE, IF YOUR DATA IS COLLECTED IN QUARTERS, YOU CAN LOOK",/,
1800  "             AT THE PREDICTIVE ABILITY OF BOTH MODELS FOR AN 18 MONTH TIME SPAN BY SPECIFYING",/,
1810  "             '6'.  AFTER ALL PREDICTIVE ABILITY TEST SITUATIONS ARE PRINTED, THE PROGRAM",/,
1820  "             PRINTS A SUMMARY OF THE TEST RESULTS.",//,
1830  "       NO  - SUPPRESSES THIS OPTION.",//,
1840  "       3.  DO YOU WANT PROJECTION AND SENSITIVITY MATRIX?",//,
1850  "       YES - WILL CAUSE PRINTING OF PROJECTION AND SENSITIVITY MATRIX.  THIS MATRIX PRESENTS",/,
1860  "             PROJECTED DIRECT LABOR REQUIREMENTS FOR SELECTED PAIRS OF CUMULATIVE PRODUCTION",/,
1870  "             PLOT POINTS AND PRODUCTION RATES.  THE PROJECTION INTERVAL FOR THE CUMULATIVE",/,
1880  "             PRODUCTION PLOT POINT IS 1% OF THE LAST OBSERVED VALUE.  THE PROJECTION VALUES",/,
1890  "             FOR PRODUCTION RATE ARE 70, 80, 90, 100, 110, 120, 130, 140, AND 150 PERCENT OF",/,
1900  "             THE LAST OBSERVED VALUE OF PRODUCTION RATE.",//,
1910  "       NO  - SUPPRESSES THIS OPTION.")
1920C
1930C
1940C*****************************************************************************************
1950C    DIMENSIONING VARIABLES
1960C*****************************************************************************************
1970       ALPHA ANSWER(10)
1980       FILENAME DATAFILE
1990       DIMENSION PLOT(150),RATE(150),HRS(150),Y(150),X1(150),X2(150),LN(150),
2000&      NEWPLOT(150),PRORATE(15),FHRS(150,15),ADEVR(999),ADEVF(999)
2010       DATA SUMHRS,SUMX1,SUMX2,SUMY,SSX1,SSX2,SUMX1Y,SUMX2Y,SMX1X2,
2020&      SSE,SSE1,SSE2,SSEL,SSEL1,SSEL2,SSTO,SSTO1,SSTO2,SSTOL,SSTOL1,SSTOL2/21*0/
2030C*****************************************************************************************
2040C
2050C    PART I - BEGIN PROGRAM, INSTRUCTIONS, DATA INPUT, DATA TRANSFORMATION, AND OPTION SELECTIONS.
2060C
2070C*****************************************************************************************
2080       PRINT,"                   THE CUMULATIVE PRODUCTION AND PRODUCTION RATE COST MODEL"
2090C:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
2100C     INSTRUCTIONS OPTION SELECTION
2110C:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
      PRINT," "
      PRINT 10
   10 FORMAT(1X,"DO YOU WANT INSTRUCTIONS?",//)
  100 INPUT, ANSWER(1)
      IF (ANSWER(1).EQ."NO") GO TO 102
      IF (ANSWER(1).EQ."YES") GO TO 101
      PRINT," ANSWER YES OR NO ONLY PLEASE"
      PRINT," "
      GO TO 100
  101 PRINT 5
      PRINT 6
      PRINT 70
C:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     INPUT THE DATA AND TRANSFORM THE VARIABLES TO LOGARITHMS
C:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C
  102 PRINT 20
   20 FORMAT(1X,"PLEASE ENTER THE NAME OF YOUR DATA FILE",//)
      INPUT, DATAFILE
      READ(DATAFILE,25)LN(1),NCASES
   25 FORMAT(V)
      DO 30 I=1,NCASES
      READ(DATAFILE,25)LN(I),HRS(I),PLOT(I),RATE(I)
      Y(I)  = ALOG10(HRS(I))
      X1(I) = ALOG10(PLOT(I))
      X2(I) = ALOG10(RATE(I))
      SUMHRS = SUMHRS + HRS(I)
      SUMX1  = SUMX1  + X1(I)
      SUMX2  = SUMX2  + X2(I)
      SUMY   = SUMY   + Y(I)
      SSX1    = SSX1    + X1(I)**2
      SSX2    = SSX2    + X2(I)**2
      SSY     = SSY + Y(I)**2
      SUMX1Y = SUMX1Y + X1(I)*Y(I)
      SUMX2Y = SUMX2Y + X2(I)*Y(I)
      SMX1X2 = SMX1X2 + X1(I)*X2(I)
   30CONTINUE
C:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     DATA CHECK OPTION SELECTION
C:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
      PRINT 35,DATAFILE
   35 FORMAT(1X,"DO YOU WANT TO CHECK DATA AS IT IS READ FROM FILE ",A8," AND CONVERTED TO L[illegible]ITHMS?",//)
  103 INPUT,ANSWER(2)
      IF (ANSWER(2).EQ."NO") GO TO 104
      IF (ANSWER(2).EQ."YES") GO TO 104
      PRINT," ANSWER YES OR NO ONLY PLEASE"
      PRINT," "
      GO TO 103
C:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     PREDICTIVE ABILITY TEST OPTION SELECTION
C:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
 104  PRINT 40
   40 FORMAT(1X,"DO YOU WANT A COMPARISON OF THE SHORTRANGE PREDICTIVE ABILITY OF THE TWO MODELS?",//)
 105  INPUT,ANSWER(3)
      IF (ANSWER(3).EQ."NO") GO TO 106
      IF (ANSWER(3).EQ."YES") GO TO 203
      PRINT," ANSWER YES OR NO ONLY PLEASE"
      PRINT," "
      GO TO 105
 203  PRINT 42
   42 FORMAT(1X,"HOW MANY CASES DO YOU WISH TO TRUNCATE?",//)
      INPUT,ITRUNC
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     PROJECTION AND SENSITIVITY MATRIX OPTION SELECTION
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 106  PRINT 45
   45 FORMAT(1X,"DO YOU WANT PROJECTION AND SENSITIVITY MATRIX?",//)
 107  INPUT,ANSWER(4)
      IF (ANSWER(4).EQ."NO") GO TO 108
      IF (ANSWER(4).EQ."YES") GO TO 108
      PRINT," ANSWER YES OR NO ONLY PLEASE"
      PRINT," "
      GO TO 107
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     BEGIN DATA CHECK OPTION
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 108  IF (ANSWER(2).EQ."NO") GO TO 109
      PRINT 50, DATAFILE
   50 FORMAT(1H1,//,75("*"),/,5X,"INPUT DATA AS READ FROM FILE ",   ," AND CONVERTED TO LOGARITHMS",
&      /,75("*"))
      PRINT," LINE    DIRECT LABOR HOURS  * CUM PROD PLOT POINT * PRODUCTION RATE"
      PRINT," NUMBER  RATIONAL  LOGARITHM * RATIONAL  LOGARITHM * RATIONAL  LOGARITHM"
      DO 60 I=1,NCASES
      PRINT 55,LN(I),HRS(I),Y(I),PLOT(I),X1(I),RATE(I),X2(I)
   55 FORMAT(1X,1X,I3,5X,F8.2,2X,F9.7," * ",F8.2,2X,F9.7," * ",F8.2,2X,F9.7)
   60CONTINUE
      PRINT 65
   65 FORMAT (1X,75("*"))
 109  PRINT 70
   70 FORMAT(1H1,15X)
C**********************************************************************************************
C
C     PART II - PEARSON CORRELATION COEFFICIENTS AND REGRESSION ANALYSIS
C
C**********************************************************************************************
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     CALCULATE AND PRINT PEARSON CORRELATION COEFFICIENTS
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
      RX1Y  = (SUMX1Y-SUMX1*SUMY/NCASES)/SQRT((SSX1-(SUMX1**2/NCASES))*(SSY-(SUMY**2/NCASES)))
      RX2Y  = (SUMX2Y-SUMX2*SUMY/NCASES)/SQRT((SSX2-(SUMX2**2/NCASES))*(SSY-(SUMY**2/NCASES)))
      RX1X2 = (SMX1X2-SUMX1*SUMX2/NCASES)/SQRT((SSX1-(SUMX1**2/NCASES))*(SSX2-(SUMX2**2/NCASES)))
      RX1X1 = 1.0
      RX2X2 = 1.0
      RYY   = 1.0
      PRINT 71,RYY,RX1Y,RX2Y,RX1Y,RX1X1,RX1X2,RX2Y,RX1X2,RX2X2
   71 FORMAT(1X,///,1X,45("*"),/,4X,"PEARSON CORRELATION COEFFICIENTS ",
&     "MATRIX",/,1X,45("*"),//,6X,"*",5X,"Y",6X,"*",5X,"X1",5X,"*",5X,
&     "X2",/,1X,45("*"),/,2X,"Y",3X,3("* ",F10.7,1X),/,1X,45("*"),/,2X,
&     "X1",2X,3("* ",F10.7,1X),/,1X,45("*"),/,2X,"X2",2X,3("* ",F10.7,1X),////)
      PRINT 70
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     CALCULATE AND PRINT THE RECRESSION RESULTS OF THE STANDARD LEARNINC CURVE MODEL
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
      B1 =(SUMX1Y-((SUMX1*SUMY)/NCASES))/(SSX1-(SUMX1**2/NCASES))
      YBAR = SUMY/NCASES
      HRSBAR = SUMHRS/NCASES
      X1BAR = SUMX1/NCASES
      X2BAR = SUMX2/NCASES
      B0 = YBAR- B1*X1BAR
      AB0 = 10.**B0
      PRINT 75
   75 FORMAT(1X,75("*"),/,14X,"RESULTS OF THE STANDARD LEARNING",
&     " CURVE MODEL",/,1X,75("*"),/,1X,"CASE",3X,"OBSERVED",5X,"PREDICTED",
&     5X,"RESIDUAL",5X,"% DEVIATION")
     DO 110 I=1,NCASES
      YHATL  = B0 + B1 * X1(I)
      RESIDL = Y(I) - YHATL
      SSEL1  = SSEL1 + RESIDL ** 2
      SSTOL1 = SSTOL1 + (Y(I) - YBAR) ** 2
      YHAT   = 10 ** YHATL
      RESID  = HRS(I) - YHAT
      PERCENT= (RESID / HRS(I)) * 100
      SSE1   = SSE1 + RESID ** 2
      SSTO1  = SSTO1 + (HRS(I) - HRSBAR) ** 2
      PRINT 80,I,HRS(I),YHAT,RESID,PERCENT
   80 FORMAT(1X,I3,4X,F8.2,6X,F8.2,5X,F8.2,7X,F6.2)
  110 CONTINUE
```

```
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     CALCULATE AND PRINT STATISTICS FOR THE STANDARD LEARNING CURVE MODEL
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
      NDFD  = NCASES - 2
      TMSRL = (SSTOL1 - SSEL1)
      TMSEL = SSEL1 / NDFD
      SEE   = SQRT (TMSEL)
      VARB0 = SEE / (1 / NCASES + X1BAR **2 / (SSX1 - (SUMX1 ** 2 / NCASES)))
      SEB0  = SQRT (VARB0)
      SEB1  = SEE / (SQRT(SSX1-(SUMX1**2/NCASES)))
      RSQL1 = (SSTOL1 - SSEL1) / SSTOL1
      RSQA1 = (SSTO1- SSE1) / SSTO1
      FRATIO= TMSRL / TMSEL
      PLEARN= (10 ** (B1 * ALOG10(2.0))) * 100
      PRINT 81,B0,SEB0,AB0,B1,SEB1,RSQL1,SEE,TMSEL,TMSRL,FRATIO,NDFD,RSQA1,PLEARN
   81 FORMAT(1X,75("*"),/,1X,"THE EQUATION FOR THIS MODEL IS: ",
     &"      YHAT = B0 * X1 ** B1",/,1X,
     &"IN LOG FORM THIS MODEL BECOMES: LOG(YHAT) = LOG(B0) + B1 * LOG(X1)",
     &/,1X,"WHERE: LOG(B0) =",F8.5,4X,"STD ERROR =",F8.5,4X,"B0 =",F11.5,
     &/,13X,"B1 =",F8.5,4X,"STD ERROR =",F8.5,
     &/,1X,"SUMMARY STATISTICS:",/,1X,
     &"R SQUARED LOG    =",F7.5,10X,"STD ERROR EST =",F11.4,/,1X,
     &"MSE",13X,"=",F9.5,8X,"MSR",11X,"=",F9.5,/,1X,
     &"F RATIO",9X,"=",F9.4,8X,"D. F. (N/D)    = 1/",I3,/1X,
     &"R SQUARED ACTUAL=",F7.5,8X,"LEARNING FACTOR =",F9.5," PERCENT",
     &/,1X,75("*"))
      PRINT 70
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     CALCULATE AND PRINT THE REGRESSION RESULTS FOR THE REDUCED HRS VS RATE MODEL
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
      B2 =(SUMX2Y-((SUMX2*SUMY)/NCASES))/(SSX2-(SUMX2**2/NCASES))
      B0 = YBAR - B2 * X2BAR
      AB0= 10 ** B0
      PRINT 82
   82 FORMAT(1X,75("*"),/,11X,"RESULTS OF REGRESSION ON PRODUCTION",
     &" RATE VARIABLE ALONE",/,1X,75("*"),/,1X,"CASE",3X,"OBSERVED",5X,
     &"PREDICTED",5X,"RESIDUAL",5X,"% DEVIATION")
      DO 111 I=1,NCASES
      YHATL  = B0 + B2 * X2(I)
      RESIDL = Y(I) - YHATL
      SSEL2   = SSEL2 + RESIDL **2
      SSTOL2  = SSTOL2 + (Y(I) - YBAR) ** 2
      YHAT    = 10 ** YHATL
      RESID  = HRS(I) - YHAT
      PERCENT= (RESID / HRS(I)) * 100
      SSE2   = SSE2 + RESID **2
      SSTO2  = SSTO2 + (HRS(I) - HRSBAR) ** 2
      PRINT 83,I,HRS(I),YHAT,RESID,PERCENT
  111 CONTINUE
```

```
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     CALCULATE AND PRINT STATISTICS FOR THE REDUCED HRS VS RATE MODEL
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
      TMSRL = (SSTOL2 - SSEL2)
      TMSEL = SSEL2 / NDFD
      SEE   = SQRT(TMSEL)
      VARB0 = SEE / (1 / NCASES + X2BAR ** 2 / (SSX2 - (SUMX2 ** 2 / NCASES)))
      SEB0  = SQRT (VARB0)
      SEB2  = SEE / (SQRT (SSX2 - (SUMX2 ** 2 / NCASES)))
      RSQL2 = (SSTOL2 - SSEL2) / SSTOL2
      RSQA2 = (SSTO2- SSE2) / SSTO2
      FRATIO= TMSRL / TMSEL
      PRINT 83,B0,SEB0,AB0,B2,SEB2,RSQL2,SEE,TMSEL,TMSRL,FRATIO,NDFD,RSQA2
   83 FORMAT(1X,75("*"),/,1X,"THE EQUATION FOR THIS MODEL IS: ",
&     "       YHAT = B0 * X2 ** B2",/,1X
&     "IN LOG FORM THIS MODEL BECOMES: LOG(YHAT) = LOG(B0) + B2 * LOG(X2)",
&     /,1X,"WHERE: LOG(B0) =",F8.5,4X,"STD ERROR =",F8.5,4X,"B0 =",F11.5,
&     /,13X,"B2 =",F8.5,4X,"STD ERROR =",F8.5,
&     /,1X,"SUMMARY STATISTICS:",/,1X,
&     "R SQUARED LOG    =",F7.5,10X,"STD ERROR EST =",F11.4,/,1X,
&     "MSE",13X,"=",F9.5,8X,"MSR",11X,"=",F9.5,/,1X,
&     "F RATIO",9X,"=",F9.4,8X,"D. F. (N/D)    = 1/",I3,/,1X,
&     "R SQUARED ACTUAL=",F7.5,/,1X,75("*"))
      PRINT 70
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     CALCULATE AND PRINT THE REGRESSION RESULTS FOR THE FULL MODEL
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
      DENOM = ((SSX1-X1BAR*SUMX1)*(SSX2-X2BAR*SUMX2) - (SMX1X2-X1BAR*SUMX2)**2)
      B1 =((SSX2-X2BAR*SUMX2)*(SUMX1Y-X1BAR*SUMY) -
&         (SMX1X2-X1BAR*SUMX2)*(SUMX2Y-X2BAR*SUMY))/DENOM
      B2 = ((SSX1-X1BAR*SUMX1) *(SUMX2Y-X2BAR*SUMY) -
&        (SMX1X2-X1BAR*SUMX2)*(SUMX1Y-X1BAR*SUMY))/DENOM
      B0 = YBAR-B1*X1BAR-B2*X2BAR
      AB0 = 10.**B0
      PRINT 84
   84 FORMAT(1X,75("*"),/,6X,"RESULTS OF COMBINED CUMULATIVE PRODUCTION",
&     " AND PRODUCTION RATE MODEL",/,1X,75("*"),/,1X,"CASE",3X,"OBSERVED",5X,
&     "PREDICTED",5X,"RESIDUAL",5X,"% DEVIATION")
     DO 112 I=1,NCASES
      YHATL = B0 + B1 * X1(I) + B2 * X2(I)
      RESIDL= Y(I) - YHATL
      SSEL  = SSEL + RESIDL ** 2
      SSTOL = SSTOL + (Y(I) - YBAR) ** 2
      YHAT  = 10 ** YHATL
      RESID = HRS(I) - YHAT
      PERCENT= (RESID / HRS(I)) * 100
      SSE   = SSE + RESID ** 2
      SSTO  = SSTO + (HRS(I) - HRSBAR) ** 2
      PRINT 85,I,HRS(I),YHAT,RESID,PERCENT
  112 CONTINUE
```

```
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     CALCULATE AND PRINT STATISTICS FOR THE FULL MODEL
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
      NDFD  = NCASES - 3
      TMSRL = (SSTOL - SSEL) / 2
      TMSEL = SSEL / NDFD
      SEE   = SQRT(TMSEL)
      ZVAL  = NCASES*(SSX1 * SSX2 - SMX1X2 ** 2) - SUMX1*(SUMX1 * SSX2 -
&             SMX1X2 * SUMX2) + SUMX2*(SUMX1 * SMX1X2 - SSX1 * SUMX2)
      AVAL  = (SSX1 * SSX2 - SMX1X2 ** 2) / ZVAL
      VARBØ = TMSEL * AVAL
      SEBØ  = SQRT (VARBØ)
      SEB1  = SQRT((TMSEL * (SSX2 - X2BAR * SUMX2)) / DENOM)
      SEB2  = SQRT((TMSEL * (SSX1 - X1BAR * SUMX1)) / DENOM)
      RSQL  = (SSTOL - SSEL) / SSTOL
      RSQA  = (SSTO - SSE) / SSTO
      FRATIO= TMSRL /TMSEL
      FB1   = (RSQL - RSQL2) / ((1 - RSQL) / (NCASES - 3))
      FB2   = (RSQL - RSQL1) / ((1 - RSQL) / (NCASES - 3))
      PRINT 85,BØ,SEBØ,ABØ,B1,SEB1,FB1,B2,SEB2,FB2,RSQL,SEE,TMSEL,TMSRL,FRATIO,NDFD,RSQA
   85 FORMAT(1X,75("*"),/,1X,"THE EQUATION FOR THIS MODEL IS: ",
&     "     YHAT = BØ * X1 ** B1 * X2 **B2",/,1X,
&     "IN LOG FORM THIS MODEL BECOMES: LOG(YHAT) = LOG(BØ) + B1 * LOG(X1) + B2 * LOG(X2)",
&     /,1X,"WHERE: LOG(BØ) =",F8.5,4X,"STD ERROR =",F8.5,4X,"BØ =",F11.5,
&     /,13X,"B1 =",F8.5,4X,"STD ERROR =",F8.5,4X,"F* =",F10.4,/,
&     13X,"B2 =",F8.5,4X,"STD ERROR =",F8.5,4X,"F* =",F10.4,/,1X,
&     "SUMMARY STATISTICS:",/,1X,"R SQUARED LOG    =",F7.5,10X,
&     "STD ERROR EST =",F11.4,/,1X,"MSE",13X,"=",F9.5,8X,"MSR",11X,"=",F9.5,/,1X,
&     "F RATIO",9X,"=",F9.4,8X,"D. F. (N/D)    = 2/",I3,/,1X,
&     "R SQUARED ACTUAL=",F7.5,/,1X,75("*"))
      PRINT 70
C****************************************************************************************************
C
C     PART III - PREDICTIVE ABILITY TEST OPTION
C
C****************************************************************************************************
      IF (ANSWER(3).EQ."NO") GO TO 116
      DO 113 I=1,ITRUNC
      ITEST = NCASES + 1 - I
      PRINT 86,ITEST,HRS(ITEST)
   86 FORMAT(1X,116("*"),/,1X,"*",37X,"SHORTRANGE PREDICTIVE ABILITY ",
&     "COMPARISON",37X,"*",/,1X,"*",16X,
&     "THE DATA PRESENTED BELOW IS FOR CASE #",I3," WHICH HAS AN OBSERVED",
&     " VALUE OF:",F9.2,16X,"*",/,
&     1X,116("*"),/1X,"*",3X,"#",3X,"*",9X,"REDUCED (LEARNING CURVE) ",
&     "MODEL",8X,"**",3X,"FULL (CUMULATIVE PRODUCTION & PRODUCTION RATE) ",
&     "MODEL",2X,"*",/,1X,"*"," CASES ",108("*"),/,1X,"* USED * ",
&     "PREDICTION * % DEVIATION *  EST BØ  * EST B1  ** ",
&     "PREDICTION * % DEVIATION *  EST BØ  * EST B1  * EST B2  *",
&     /,1X,116("*"))
      DO 114 J=1,ITRUNC
```

```
      ICASES = ITEST - J
      SUMY   = 0
      SUMX1  = 0
      SUMX2  = 0
      SSX1   = 0
      SSX2   = 0
      SUMX1Y = 0
      SUMX2Y = 0
      SMX1X2 = 0
     DO 115 K=1,ICASES
      SUMY   = SUMY + Y(K)
      SUMX1  = SUMX1 + X1(K)
      SUMX2  = SUMX2 + X2(K)
      SSX1   = SSX1 + X1(K) ** 2
      SSX2   = SSX2 + X2(K) ** 2
      SUMX1Y = SUMX1Y + X1(K) * Y(K)
      SUMX2Y = SUMX2Y + X2(K) * Y(K)
      SMX1X2 = SMX1X2 + X1(K) * X2(K)
 115 CONTINUE
      ICOUNTA = ICOUNTA + 1
      YBAR  = SUMY / ICASES
      X1BAR = SUMX1 / ICASES
      X2BAR = SUMX2 / ICASES
      B1R   = (SUMX1Y - ((SUMX1 * SUMY) / ICASES)) / (SSX1 - (SUMX1 ** 2 / ICASES))
      B0R   = YBAR - B1R * X1BAR
      AB0R  = 10 ** B0R
      YHATR = 10 ** (B0R + B1R * X1(ITEST))
      DEVR    = HRS(ITEST) - YHATR
      ADEVR(ICOUNTA) = ABS(DEVR)
      SUMADEVR = SUMADEVR + ADEVR(ICOUNTA)
      PDEVR   = 100 * DEVR/HRS(ITEST)
      APDEVR = ABS(PDEVR)
      IF (APDEVR.GT.10.0) GO TO 201
      ICOUNTGR = ICOUNTGR + 1
      IF (APDEVR.GT.5.0) GO TO 201
      ICOUNTER = ICOUNTER + 1
 201  DENOM = ((SSX1-X1BAR*SUMX1)*(SSX2-X2BAR*SUMX2) - (SMX1X2-X1BAR*SUMX2)**2)
      B1F   = ((SSX2-X2BAR*SUMX2)*(SUMX1Y-X1BAR*SUMY) -
&             (SMX1X2-X1BAR*SUMX2)*(SUMX2Y-X2BAR*SUMY))/DENOM
      B2F   = ((SSX1-X1BAR*SUMX1)*(SUMX2Y-X2BAR*SUMY) -
&             (SMX1X2-X1BAR*SUMX2)*(SUMX1Y-X1BAR*SUMY))/DENOM
      B0F   = YBAR - B1F * X1BAR - B2F * X2BAR
      AB0F  = 10 ** B0F
      YHATF = 10  ** (B0F + B1F * X1(ITEST) + B2F * X2(ITEST))
      DEVF  = HRS(ITEST) - YHATF
      ADEVF(ICOUNTA) = ABS(DEVF)
      SUMADEVF = SUMADEVF + ADEVF(ICOUNTA)
      PDEVF = 100 * DEVF/HRS(ITEST)
      APDEVF = ABS(PDEVF)
      IF (APDEVF.GT.10.0) GO TO 202
      ICOUNTGF = ICOUNTGF + 1
      IF (APDEVF.GT.5.0) GO TO 202
      ICOUNTEF = ICOUNTEF + 1
 202  PRINT 87,ICASES,YHATR,PDEVR,AB0R,B1R,YHATF,PDEVF,AB0F,B1F,B2F
   87 FORMAT(1X,"*",2X,I3,2X,"*",1X,F9.2,2X,"*",3X,F6.2,4X,"*",F9.2,1X,
&     "*",F8.5,1X,"**",1X,F9.2,2X,"*",3X,F6.2,4X,"*",F9.2,1X,"*",F8.5,1X,
&     "*",F8.5,1X,"*")
```

```
  114 CONTINUE
      PRINT 88
   88 FORMAT(1X,116("*"),//////)
      COUNT = COUNT + 1.0
      FLAG1 = COUNT / 2.0
      FLAG2 = FLAG1 - INT(FLAG1)
      IF (FLAG2.NE.0.0) GO TO 113
      PRINT 79
  113 CONTINUE
      AVGADEVR = SUMADEVR / ICOUNTA
      AVGADEVF = SUMADEVF / ICOUNTA
     DO 119 I =1,ICOUNTA
      SSDEVR  = SSDEVR + (ADEVR(I) - AVGADEVR)**2
      SSDEVF  = SSDEVF + (ADEVF(I) - AVGADEVF)**2
  119 CONTINUE
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
C     CALCULATE AND PRINT RESULTS SUMMARY FOR PREDICTIVE ABILITY TESTS
C::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
      VARADEVR = SSDEVR / (ICOUNTA - 1)
      VARADEVF = SSDEVF / (ICOUNTA - 1)
      TESTSTAT = (AVGADEVR-AVGADEVF)/SQRT((VARADEVR/ICOUNTA)+(VARADEVF/ICOUNTA))
      PCENTER = 100 * ICOUNTER / ICOUNTA
      PCENTCR = 100 * ICOUNTCR / ICOUNTA
      PCENTEF = 100 * ICOUNTEF / ICOUNTA
      PCENTCF = 100 * ICOUNTCF / ICOUNTA
      PRINT 95,AVGADEVR,AVGADEVF,VARADEVR,VARADEVF,TESTSTAT,ICOUNTA,
&     ICOUNTA,ICOUNTER,ICOUNTEF,PCENTER,PCENTEF,ICOUNTCR,ICOUNTCF,PCENTCR,
&     PCENTCF
   95 FORMAT(1X,67("*"),/,1X,"*",10X,"SUMMARY OF PREDICTIVE ABILITY TESTS",
&     " RESULTS",12X,"*",/,1X,67("*"),/,1X,"*",9X,"ITEMS OF INTEREST",8X,
&     "* REDUCED MODEL *  FULL MODEL  *",/,1X,67("*"),/,1X,"* AVERAGE ",
&     "ABSOLUTE DEVIATION",7X,"*",3X,F9.2,3X,"*",2X,F9.2,3X,"*",/,1X,
&     "* VARIANCE OF ABSOLUTE DEVIATIONS",2X,"*",1X,F11.2,3X,"*",F11.2,3X,
&     "*",/,1X,"* TEST STATISTIC (SEE NOTE)",8X,"*",6X,"---",6X,"*",2X,
&     F9.2,3X,"*",/,1X,"* TOTAL NUMBER OF TEST SITUATIONS  *",6X,I3,6X,"*",
&     5X,I3,6X,"*",/,1X,"* NUMBER OF PREDICTIONS WITHIN 5%  *",6X,I3,6X,
&     "*",5X,I3,6X,"*",/,1X,"* PERCENT OF PREDICTIONS WITHIN 5% *",6X,F4.0,
&     5X,"*",5X,F4.0,5X,"*",/,1X,"* NUMBER OF PREDICTIONS WITHIN 10% *",
&     6X,I3,6X,"*",5X,I3,6X,"*",/,1X,"* PERCENT OF PREDICTIONS WITHIN 10%*",
&     6X,F4.0,5X,"*",5X,F4.0,5X,"*",/,1X,67("*"),/,1X,"NOTE: IN TESTING FOR ",
&     "STATISTICAL SIGNIFICANCE USE STUDENT'S T DISTRIBUTION",/,1X
&     "IF THE NUMBER OF TEST SITUATIONS ARE LESS THAN 60; OTHERWISE ",
&     "USE STANDARD",/,1X,"NORMAL DISTRIBUTION.  IN EITHER CASE THIS IS ",
&     "A ONE TAILED TEST.  IF",/,1X,"THE TEST STATISTIC IS GREATER THAN ",
&     "THE CRITICAL STATISTIC ONE MAY",/,1X,"CONCLUDE THAT THE AVERAGE ",
&     "ABSOLUTE DEVIATION OBTAINED WITH THE FULL",/,1X,"MODEL IS ",
&     "SIGNIFICANTLY LESS THAN THAT OBTAINED WITH THE REDUCED MODEL.")
```

```
C*****************************************************************************************************
C
C     PART IV - PROJECTION AND SENSITIVITY MATRIX OPTION
C
C*****************************************************************************************************
  116 IF (ANSWER(4).EQ."NO") GO TO 125
      ADDPLOT = PLOT(NCASES)
      PRINT 70
     DO 117 I=1,100
      ADDPLOT = ADDPLOT + .01 * PLOT(NCASES)
      NEWPLOT(I) = INT(ADDPLOT)
      ADDRATE = .60 * RATE(NCASES)
     DO 118 J=1,9
      ADDRATE = ADDRATE + .1 * RATE(NCASES)
      PRORATE(J) = ADDRATE
      FHRS(I,J)  = AB0 * NEWPLOT(I)**B1 * PRORATE(J)**B2
  118 CONTINUE
  117 CONTINUE
      ISTART = 1
      ISTOP  = 50
     DO 120 K=1,2
      PRINT 89,(PRORATE(J),J=1,9)
   89 FORMAT(1X,113("*"),/,1X,"*",39X,"PROJECTION AND SENSITIVITY MATRIX",
&     39X,"*",/,1X,113("*"),/,1X,"*  PROJECTED *",36X,"PROJECTED PRODUCTION",
&     " RATES",36X,"*",/,1X,"* CUMULATIVE *",99("*"),/,1X,"*    UNITS    *",
&     9(F8.2,2X,"*"),/,1X,113("*"))
     DO 121 I=ISTART,ISTOP
      PRINT 90,NEWPLOT(I),(FHRS(I,J),J=1,9)
   90 FORMAT(1X,"*",3X,I6,3X,"*",9(1X,F8.1,1X,"*"))
  121 CONTINUE
      PRINT 91
   91 FORMAT(1X,113("*"))
      PRINT92
   92 FORMAT(1X,"NOTE: 1.  PROJECTED VALUES FOR DIRECT LABOR HOURS MAY ",
&     "BE READ FROM THE ABOVE MATRIX BY MATCHING A GIVEN PRODUCTION",/,1X,
&     "RATE WITH A GIVEN NUMBER OF CUMULATIVE UNITS AND READING THE ",
&     "VALUE FOR DIRECT LABOR HOURS FOUND AT THE INTERSECTION",/,1X,
&     "OF THE CORESPONDING ROW AND COLUMN.  FORECASTING MODEL IS THE ",
&     "CUMULATIVE PRODUCTION & PRODUCTION RATE MODEL.",/,7X,"2.  PROJECT",
&     "ION INTERVAL FOR CUMULATIVE UNITS IS 1% OF THE LAST OBSERVED VALUE",
&     " OF CUMULATIVE UNITS.",/,7X,"3.  PROJECTION VALUES FOR PRODUCTION ",
&     "RATE ARE 70, 80, 90, 100, 110, 120, 130, 140, AND 150 PERCENT OF ",
&     "THE",/,1X,"LAST OBSERVED VALUE OF PRODUCTION RATE.")
      ISTART = 51
      ISTOP  = 100
      PRINT 70
  120 CONTINUE
  125 STOP
      END
```

SELECTED BIBLIOGRAPHY

A. REFERENCES CITED

1. Alchian, Armen A., and William R. Allen. University Economics. Belmont, California: Wadsworth Publishing Company, Inc., 1964.

2. Clark, Donald S., and Thomas F. McNeill. Cost Estimating and Contract Pricing. New York: American Elsevier Publishing Company, Inc., 1966.

3. Congleton, Captain Duane E., USAF, and Major David W. Kinton, USAF. "An Empirical Study of the Impact of a Production Rate Change on the Direct Labor Requirements for an Airframe Manufacturing Program." Unpublished master's thesis, LSSR 23-77B, AFIT/LSGR, Wright-Patterson AFB, OH 1977. AD-A-52720.

4. Dreyfuss, David J., and Joseph P. Large. "Estimated Costs of Extended Low-Rate Airframe Production." Unpublished research report No. R-2243-AF, The Rand Corporation, Santa Monica, CA, 1978.

5. Ilderton, Robert Blair. "Methods of Fitting Learning Curves to Lot Data Based on Assumptions and Techniques of Regression Analysis." Unpublished master's thesis, George Washington University, Washington, D.C., 1970. AD-A011583.

6. Johnson, Gordon J. "The Analysis of Direct Labor Costs for Production Program Stretchouts," National Management Journal, Spring 1969, pp. 25-41.

7. Large, Joseph P., Karl Hoffmayer, and Frank Kontrovich. "Production Rate and Production Cost." Unpublished research report No. R-1609-PA&E, The Rand Corporation, Santa Monica, CA, 1974.

8. Neter, John, and William Wasserman. Applied Linear Statistical Models. Homewood, IL: Richard D. Irwin, Inc., 1974.

9. Orsini, Captain Joseph A., USAF. "An Analysis of Theoretical and Empirical Advances in Learning Curve Concepts Since 1966." Unpublished master's thesis, GSA/SM/70-12, AFIT/SE, Wright-Patterson AFB, OH, 1970. AD-875892.

10. Ostwald, Phillip F. Cost Estimating for Engineering and Management. Englewood Cliffs, N.J.: Prentice-Hall Inc., 1974.

11. Smith, Lieutenant Colonel Larry L., USAF. "An Investigation of Changes in Direct Labor Requirements Resulting from Changes in Airframe Production Rate." Unpublished doctoral dissertation, Department of Marketing, Transportation and Business Environment, University of Oregon, Eugene, Oregon, 1976. AD-A926112.

12. U. S. Department of Defense. Annual Report Fiscal Year 1979. Washington, D.C.: Government Printing Office, 1978.

13. ________. Armed Services Procurement Regulation Manual (ASPM No. 1). Washington, D.C.: Government Printing Office, September 1975.

14. Wonnacott, Thomas H. and Ronald J. Wonnacott. Introductory Statistics for Business and Economics. 2nd ed. New York: John Wiley and Sons, 1977.

B. RELATED SOURCES

Brewer, Glenn M. "The Learning Curve in the Airframe Industry." Unpublished master's thesis, SLSR-18-65, AFIT/SL, Wright-Patterson AFB, OH, 1965.

Brockman, Major William F., USAF, and Major Freddie D. Dickens, USAF. "Investigation of Learning Curve and Cost Estimation Methods for Cargo Aircraft." Unpublished research paper, SGM/SM/67-2/7, AFIT/SE, Wright-Patterson AFB, OH, 1967. AD-665464.

BIOGRAPHICAL SKETCHES

BIOGRAPHICAL SKETCHES

Captain Michael W. Crozier graduated from Missouri Valley College in June 1971 with a Bachelor of Science degree in Mathematics. Upon commissioning through the Officer Training School in August 1972, he entered Undergraduate Navigator Training at Mather AFB, CA. He was then assigned to the 68th Bomb Wing at Seymour-Johnson AFB, NC as a B-52 navigator. Prior to AFIT he was a B-52 radar navigator at Seymour-Johnson. Upon graduation from the School of Systems and Logistics he will be assigned to the Contracting Branch, Electronics System Division of AFSC at Hanscom Field, MA.

Captain Edward J. J. McGann, Jr. graduated from the USAF Academy in June 1971 with a Bachelor of Science Degree in Engineering Management. His first commissioned assignment was as base procurement officer at Offutt AFB, NB. In April 1973 he entered Undergraduate Navigator Training at Mather AFB, CA. He was then assigned to the 374 TAW in the West Pacific, where he performed duties as C-130 navigator, instructor navigator, and wing training officer. Upon graduation from the School of Systems and Logistics, he will enter the AFLC Logistics Career Broadening Program at the San Antonio ALC.